UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA

UNIVERSITÀ DEGLI STUDI
DI MODENA E REGGIO EMILIA

Dottorato di ricerca in
Scienze Umanistiche

Ciclo XXXVIII

MORAL AND POLITICAL CONSTRUCTIVISM IN THE HUMAN RIGHTS DEBATE

Candidato: Jacopo Morelli

Relatore (Tutor): Prof. Carlo Altini

Correlatore (Co-Tutor): Prof.ssa Carla Bagnoli

Coordinatrice del Corso di Dottorato: Prof.ssa Laura Gavioli

TABLE OF CONTENTS

PREFACE **p.6**

NOTE ON CITATIONS AND ABBREVIATIONS **p.9**

INTRODUCTION **p.10**

A) The Constructivist Framework p.14

B) Moral, Ethical, Legal and Political Contexts p.17

C) Moral and Political Constructivism p.28

D) Human Rights at the Intersection of Justice Contexts p.32

E) Laying Out the Blueprint: Project Overview p.34

PART I CONCEPTUAL ANALYSIS: TOWARD A CONSTRUCTIVIST DEFINITION OF HUMAN RIGHTS **p.38**

CHAPTER 1 THE CONCEPT OF RIGHT **p.42**

1.1 The Nature: Overlapping Boundaries **p.42**

1.1.1 The Separation Thesis of Moral and Legal Rights p.42

1.1.2 Juspositivism and The Rule of Recognition p.45

1.1.3 Dworkin's Constructivist Conception of Law as Integrity p.50

1.2 The Form: Claimability and Correlativity **p.55**

1.2.1 Claim and Liberty rights p.55

1.2.2 Correlativity of duties and rights p.60

1.2.3 The Critique of O'Neill's Focus on Duties p.68

1.3 The Function: Specificatory Demand **p.76**

1.3.1 Will and Interest Theory p.76

1.3.2 Darwall and the Demand Theory of Right p.85

1.3.3 Richardson's Specificatory Theory p.90

1.4 A Preliminary Constructivist Definition of Right **p.97**

CHAPTER 2 FURTHER NORMATIVE CONCEPTS **p.102**

2.1 Normativity and Agency **p.104**

2.2 Reasoning about Reasons **p.109**

2.2.1 Motivation and Justification: Internalism Vs Externalism p.110

2.2.2 Subjective, Objective and Intersubjective Reasons p.116

2.2.3 Practical Reasoning as Instrumental, Reflexive and Transformative p.123

2.2.4 Moral Reasoning between Universalizability and Defeasibility p.129

2.3 The Value of Values **p.140**

2.3.1 The Fact-Value Distinctions: Naturalism Vs Non-Naturalism p.143

2.3.2 Moral Value: Intrinsic or Extrinsic? p.152

2.3.3 Final ends, Value Pluralism and Incommensurability p.156

2.4 Back to Rights, and Beyond **p.162**

CHAPTER 3 THE IDEA OF HUMAN BEING **p.167**

3.1 The Concept of Dignity in The Human Right Debate **p.169**

3.1.1 Eliminating Dignity: Macklin, Pinker and Singer p.171

3.1.2 Preserving Dignity: Waldron's Reaffirmation of a Status-Concept p.173

3.1.3 Reconceptualizing Dignity: Constructivists Redefinitions p.176

3.2 Some Kantian Perspectives on Dignity **p.180**

3.2.1 A Constructivist Interpretation of Kant: Sensen p.181

3.2.2 A Realist Approach to Kantian Dignity: Caranti p.188

3.2.3 Towards a Kantian Convergence on a Relational Status p.194

3.3 Alternative Foundations **p.197**

3.3.1 The Reductive View: Human Capacity and Needs p.198

3.3.2 The Orthodox View: Human Worth and Status p.202

3.3.3 A Relational Foundation Beyond Instrumentalism and Non-Instrumentalism p.205

3.4 From Human Rights to Person Rights? p.208

PART II HISTORICAL RECONSTRUCTION: TOWARD A CONSTRUCTIVIST CONCEPTION OF UNIVERSALITY FOR HUMAN RIGHTS p.218

CHAPTER 4 CONSTRUCTIVISM AS AN ALTERNATIVE TO JUSNATURALISM AND JUSPOSITIVISM p.222

4.1 An Historical Revisitation of a Perennial Debate p.222

4.1.1 The old idea of Natural Law p.222

4.1.2 A modern conceptualization of jusnaturalism p.225

4.1.3 Kant as the Bridge Between Jusnaturalism and Juspositivism p.235

4.1.4 The contemporary conception of Juspositivism p.254

4.2 Enduring Problems p.259

4.2.1 Limitations of Jusnaturalism: Foundationalism and Arbitrariness p.259

4.2.2 Weaknesses of Juspositivism: Circularity and Formalism p.267

4.3 Moving Beyond the Resulting Theoretical Stalemate through Universality p.277

CHAPTER 5 GROUNDING UNIVERSALITY IN CONSTRUCTIVIST TERMS p.281

5.1 Early Constructivist Conceptions of Universality p.283

5.1.1 Rawls's Original Position and Reflective Equilibrium p.284

5.1.2 Habermas's Discourse Theory and Principle (U) p.296

5.1.3 Shared Strengths and Limitations of Rawls and Habermas p.305

5.2 Emerging Developments in Constructivist Universality p.312

5.2.1 O'Neill's Reasonable Universalism p.312

5.2.2 Forst's Reflexive Universality p.327

5.3. Potentials and Limits of Constructivist Universality p.337

5.3.1 Is Constructivist Universality Foundational or Formal? Inescapability of Agency p.338

5.3.2 Is Constructivist Universality Circular or Arbitrary? Transparency of Normativity p.347

5.4 The Universality of Human Rights in Constructivist Terms p.353

CONCLUSIONS: HUMAN RIGHTS AS UNIVERSAL DEMANDS p.357

REFERENCES p.373

PREFACE

This work is an attempt to explore whether it is possible to bring order to the chaotic philosophical debate on human rights through a specific approach: constructivism. It is a natural continuation of my master's thesis conducted in Parma, which focused on the constructivist approach in the moral domain. Even during the writing of that thesis, I felt that the separation of moral and political fields was often arbitrary and not well-defined; after all, John Rawls, who was one of the central figures in my master's thesis on metaethics, is primarily remembered as a political philosopher. During the transition from my master's to my doctoral studies, an in-depth reading of Jürgen Habermas and his communicative ethics as the essence of the democratic political ideal, further convinced me of the need to study the moral-political connection more closely. I believe that there is no more suitable theme for this purpose than human rights.

During the writing process, I faced the challenge of navigating the classic pitfalls of a doctoral thesis. I aimed to balance the tendency to produce an "endless treatise" against the risk of adopting an overly generalist approach; additionally, I sought to avoid a literature review lacking originality while also steering clear of an excess of personal ideas of my "own" version of constructivism. My goal was to strike a balance that I hope is evident in the final result. This task was particularly challenging since, as Shapiro noted, "philosophical debates are hard to characterize because, unlike formal debates, they are not usually about just one issue. In philosophy, everything is ultimately connected to everything else, and hence philosophical controversies tend to range over many different, though in-the-end related, questions." (Shapiro, 2007, p. 4).

In these final days of refining the whole text, I want to express my deep gratitude to all the people who contributed to enriching this journey. First, I thank the professors who have followed me from the beginning, Carlo Altini and Carla Bagnoli, for all the advice, from the most practical to the most substantive, which they have provided me over these three years. As I mentioned, this project stems from the challenge of trying to bring the political and moral domains onto the shared plane of human rights; in this sense, I could not have been more fortunate than to be supervised by a political philosopher and a moral philosopher.

I am equally grateful to Professor Rainer Forst for welcoming me to Frankfurt and for all the discussions we had about my project, from meetings in his office to dinners in the evening. It has been an honor and a privilege for me to be able to research under his expert and friendly supervision. I also wish to thank Professor Luigi Caranti for allowing me to attend his master classes on human rights in Frankfurt, which offered valuable opportunities to refine the arguments I was working on while also discovering new perspectives.

I am likewise indebted to Professor Luc Foisneau for welcoming me to Paris. It was a privilege to experience the richness of Parisian intellectual life, from the inspiring workshops organized with Nadia Urbinati and Cristina Lafont, to the "Lectio Magistralis" by Michael Sandel, as well as many other stimulating conferences and discussions.

I also wish to acknowledge my colleagues from the 38th doctoral cycle in Modena. It is never guaranteed to find peers one truly connects with, especially when research interests diverge so widely, yet I am proud of the conference on identity we succeeded in organizing by bringing together our different perspectives. I am particularly thankful to my academic siblings, Francesco Emanuelli and Manuel Marsico, for sharing this journey with me, even if we started the PhD in different years. Being able to discuss constructivism whenever a specific doubt arose—often with a quick message—proved immensely valuable, and every conference we attended together offered an opportunity to grow, whether by commenting on talks or preparing papers together as discussants.

A nostalgic thanks goes to the new friends I met in Frankfurt, like the members and guests of the Research Network Normative Orders, whose Tuesday discussions, both at the university and in the Italo-Indian pizzeria, have been an inexhaustible source of ideas and reflections for this project. I am especially indebted to Robin Waldenburg and Julio Caceda, who not only engaged in many stimulating discussions with me but also read and commented on several drafts of my work, and with whom I had the pleasure of continuing our exchanges at subsequent conferences. My sincere thanks also go to Daniel Piñeiro Rodriguez, with whom I shared countless conversations during lunches and dinners at our Gästehaus canteen, and whom I later had the pleasure of meeting again at a conference in Brazil, where we also began planning a book that we are publishing together. Collaborating with him, as well as exchanging ideas with other

postdoctoral researchers in law such as Lucas Reckziegel Weschenfelder and Pratyush Kumar, has been of immense value and has greatly enriched this work.

A heartfelt thanks goes also to my lifelong friends - Diego, Francesco, Luca and Filippo - for their stimulating discussions. Even without being experts in the field, they showed genuine interest and a sharp critical sense, which greatly enriched my work. I especially thank Ruggero Orilia, with whom the bond formed during our master's course led us to develop parallel research projects, although from different perspectives (rights/needs). Every discussion, from the tables of the Parma canteen to the beaches of Porto Recanati, represented an invaluable contribution that has enriched this academic journey immensely.

It would, of course, be impossible to name all the others I met along the way with whom I had valuable discussions during the many conferences and workshops that shaped this project. Still, I wish to at least briefly acknowledge those professors and young scholars who generously read part of my work and offered comments and insights, including Ian Carter, Joel Thiago Klein, Gabriele Badano, Wolfgang Huemer, Francesco Maiolo, Ciara Luxton, Stefano Pinzan, Luca Hemmerich, Marzia Marconi, Andrei Bespalov, Edoardo Vignocchi, Joao Freitas Mendas and Massimo Caon.

Above all, my deepest gratitude is reserved for my family and for Paula. Their patience, encouragement, and unwavering support have sustained me throughout this long journey, and this work would not have been possible without them.

In conclusion, I also feel compelled to address the current global situation. I began working on this project just a few months after the start of the Ukraine-Russia war and I was in the midst of my writing when the atrocities of the Israel-Palestine conflict unfolded. This project is divided into two parts: one primarily theoretical and the other more historical. My initial idea also included a third applied part, where I intended to utilize the constructivist framework developed in the first two parts to urgent human rights issues, such as the right of intervention. However, the writing of the first part greatly exceeded my intended expectations, leading me to stop at the reconstruction provided. This is why I feel the need to apologize for the limited focus on actual cases of human rights violations. Nevertheless, this is a project I hope to continue working on in the near future to complete and refine. Despite this, I am confident that I have at least provided a solid starting point for further research.

NOTE ON CITATIONS AND ABBREVIATIONS

The citation style for this project follows APA 7th edition guidelines, with one deliberate exception: references to Immanuel Kant's works. Given Kant's central importance for constructivist theory—both conceptually and structurally—citations of his texts follow the Akademie-Ausgabe notation: *Kant's gesammelte Schriften, herausgegeben von der Deutschen* [formerly *Königlich Preußischen*], Akademie der Wissenschaften (Berlin: de Gruyter [and predecessors], 1902. This convention is standard in Kant scholarship and is adopted here not only for reasons of disciplinary tradition but also for scholarly precision. Having consulted Kant's writings in three different languages, I found that the use of Akademie volume and page numbers enables accurate cross-referencing across multiple editions and translations. Abbreviated titles, preceding the page numbers of the relevant Akademie volume, are used as follows:

- **Anth** (*Anthropology from a Pragmatic Point of View*)
- **G** (*Groundwork for the Metaphysics of Morals*)
- **KrV (***Critique of Pure Reason***) [A** (first edition 1781**) B** (second edition 1787)**]**
- **KpV** (*Critique of Practical Reason*)
- **KU** (*Critique of Judgment*)
- **MAN** (*Metaphysical Foundations of Natural Science*)
- **MS** (*Metaphysics of Morals*)
- **P** (*Prolegomena to Any Future Metaphysics*)
- **R** (*Religion within the Limits of Bare Reason*)
- **TP (***On the Common Saying: That May Be Correct in Theory, But Is of No Use in Practice*)
- **ZeF** (*Toward Perpetual Peace*)

INTRODUCTION

"The theory of human rights is and ought to be interdisciplinary. Any adequate understanding of human rights requires contributions by lawyers and political scientists, as well as moral, political, and legal philosophers."

(Wellman, 2011b, p. 10).

"Within contexts of political discourse, the member of legal communities requires that law and morality substantively intersect."

(Habermas, 1996, p. 207).

Despite living in an age marked by profound social injustices[1] and war conflicts,[2] the human rights narrative seems so deeply rooted in our belief systems, so self-evident and acquired, that the violation of them is seen not as a failure in recognizing them, but as their willful and undue violation which can only be condemned.[3] This is not a recent tendency, considering that as early

[1] As Thomas Pogge reminded us: "About half of all human beings live in severe poverty and about a quarter live in extreme or life-threatening poverty […] About one third of all human deaths, 18 million each year, are due to poverty-related causes" (Pogge, 2014, p. 21).

[2] The war in Ukraine and Palestine is merely the closest, both chronologically and spatially. Shifting our focus to other distant regions, such as Southeast Asia (e.g., the Burma civil war and the Papua-Indonesian conflict) or West-Central Africa (notably, two coups in Gabon and Niger, alongside various merciless civil and secessionist wars like in Ambazonia and, above all, in Sudan, where it has already claimed tens of thousands of lives and displaced millions), reveals the continued prevalence of wars and violence. I do not intend to delve into the perennial discussion of whether we are experiencing more or fewer wars. Even if we align with Steven Pinker's perspective that violence has historically declined (Pinker, 2012), it remains a fact that we still live in an age marked by numerous conflicts, and this is a crucial consideration when theorizing on contemporary human rights.

[3] One could immediately object that this perception is very Western-centric and that other cultures not only do not have it but rather view human rights as a concealed tool of imperialism to dominate non-Western societies, culturally and politically. This is a central theme, which will be addressed and explored in subsequent chapters. We can here, however, already hint at what is the crucial point: even leaving aside the actual fact that every member state of the United Nations has signed one or more of the six major human rights treaties and more than 80 percent have ratified four or more of these six core documents (Nickel, 2006, p. 274), it remains conceptually possible for every person and every culture to abstract from the individual rights that have been reaffirmed historically and to think of the general concept of universal rights as legitimate claims deserving protection. Concepts similar to those of "human rights" can be found, as Jürgen Habermas points out, not only in Western culture, but also in Asian and many African

as 1964 Bobbio wrote how "the basic problem concerning human rights today is not so much that of justifying them as that of protecting them. It is not a philosophical problem, but a political one" (Bobbio, 2014, pp. 20–21). This quotation admits of several interpretations, but the most immediate one seems to indicate that the philosophical discussion on human rights (including several theoretical sub-questions, such as their foundation, legitimacy, and universality) is finished, resolved, and not further deepened, and that only the application side, relegated to politics (not as science, but as a game of powers and relations) is now incomplete. If this is the interpretation we want to give to Bobbio's quote, ignoring the underlying celebratory and optimistic intentions, there seems to be an undue oversimplification of the theoretical debate behind the topic of human rights, which is in fact still very much active and far from settled.[4]

Indeed, according to Johan Karlsson Schaffer and Reidar Maliks, the philosophy of human rights is currently embroiled in a methodological dispute between traditionalist or moral approaches, which derive universal rights directly from moral theory, and political or practical approaches, which focus on a practical analysis of international politics to derive normative claims (Maliks & Schaffer, 2017a).[5] In this sense, John Rawls's work, *The Law of People*, arguably initiated this debate by integrating considerations of how international politics actually functions

---

cultures: think of the concept of *Ren* in Confucianism, which can be partially translated as a mix of benevolence and empathy, or the South African concept of *Ubuntu*, relationship, or more specifically "being person through other people" (Habermas 1984). The idea of human rights is a concept all the cultures tend to seek. The existence of different cultures and different declinations of this concept does not imply their necessary incommensurability: most human beings belong to several cultures simultaneously. A constructivist conception of human rights wants to be precisely culture-neutral, that is, both immanent to a culture and transcendent to the culture itself, in order to achieve a basic social structure that is justified both internally and externally (Forst, 1999, 56).

[4] This statement also contrasts with another Bobbio's claim, which emphasizes the inevitably ever-changing nature of human rights by asserting that "the list of human rights has changed and continues to change with the evolving historical conditions, that is, with the changing needs and interests, the classes in power, and the means available for their implementation." (Bobbio, 2014, p. 25). This last quote also seems to indicate a legal positivist tendency, which does not align with interpretations of him as a natural law theorist. However, Bobbio is not alone in being difficult to categorize within the dichotomy of jusnaturalism and juspositivism, as I will also argue in part II.

[5] For some prominent examples of the moral conception, see Gewirth (1982), Griffin (2008), Wellman (2011b) and Cruft, Liao and Renzo (2015). For some prominent examples of the political conception, see Cohen (2004), Beitz (2009), Raz (2010a), and Buchanan (2017) . However, as Hessler (2017) herself acknowledges in the edited collection by Maliks and Schaffer, the division is not as clear-cut or oversimplified as it is often presented. There are significant divergences and variations within each approach, with scholars often distancing themselves from others grouped under the same category, like in "Raz's moral approach, which Buchanan sees as committed to the mirroring view" (Hessler, 2017, p. 28). I will revisit this division and explore how constructivism can fit within it in the conclusions of this project.

into the philosophical and normative theorizing of human rights (Rawls 1999).[6] According to Rawls, one of the defining features of human rights is that if a society fails to respect them, then external agents may in specified circumstances be permitted to interfere in its internal affairs. This interference could take various forms, such as economic or political sanctions, or even military intervention.[7]

Rawls's work has elicited contrasting reactions. On the one hand, James Griffin objects that Rawls diverts attention from the central point – namely, the foundational element that can elucidate the normative nature of these rights (Griffin 2008). On the other hand, Rawls' work garnered positive appraisals from thinkers such as Charles Beitz and Joseph Raz, who deemed the primary characteristic of human rights to be inherently political, specifically pertaining to the realm of international politics (Beitz, 2009; Raz, 2010a). In proposing that human rights should be understood in terms of their role or function in modern international political practice, the political conception directly challenges the idea that human rights include all the rights that we have simply by virtue of being human, thus creating a hiatus between their approach and the classical approach.[8] Kristen Hessler emphasized that these methodological differences reflect substantively different philosophical projects, since "political views assume (and moral ones deny) that human rights must be defined and analyzed in terms of specifically political or institutional concepts (Hessler, 2017, pp. 16–17).

In this dichotomous framing of the debate, moral reasoning within the international sphere is granted only a reductive or a distorting function: either it is tasked solely with the identification

[6] "While I believe that these forms of skepticism about natural rights are in the end indefensible—and that plausible arguments for natural rights foundations are in fact available to us—the past few decades of political philosophy/theory have been more centrally devoted to avoiding than to resolving such debates. And that same strategy of avoidance has, of course, been deployed in discussions of human rights, due largely to the influence of Rawls' account of human rights in The Law of Peoples" (Simmons, 2015, p. 146). Yet, Rawls is often regarded not only as the figure who initiated this debate but also as the one who brought it to a close. For instance, Luise Katharina Müller advocates for abandoning the division between moral and political conceptions of human rights, offering an innovative interpretation of John Rawls' integrated view of human rights (L. K. Müller, 2017).

[7] "Human rights are a class of rights that play a special role in a reasonable Law of Peoples: they restrict the justifying reasons for war and its conduct, and they specify limits to a regime's internal autonomy" (Rawls, 1999, p. 79).

[8] Mayer argues that "political approaches offer a better way of understanding the solidarity necessary for realizing human rights. Because political conceptions of solidarity determine a concrete vehicle for pursuing human rights that solidarity is built around, they are superior to moral conceptions, which fail to do so" (Mayer 2017, p.157)

and/or definition of abstract values, which are then used to evaluate whether existing international doctrines conform to them, or it is considered wholly irrelevant to legal analysis, dismissed as extraneous to the domain of international law. This debate, in essence, mirrors the historical divide between jusnaturalists and juspositivists, albeit in a different guise. In the first case, morality is treated as a transcendent normative benchmark, external to law, whose sole role is to judge legal doctrines from a superior evaluative standpoint. In the second, it is excluded from serious jurisprudential consideration, relegated to the sphere of private conviction, societal customs or even political rhetoric. One of the central arguments advanced in this project is that the widely accepted distinction between natural rights and positive rights is fundamentally misguided (as discussed in Part II)[9], and rests on an incorrect understanding of the nature and function of rights, norms, and values (as explored in Part I). By perpetuating this conceptual divide, legal theory risks entrenching an impoverished view of both law and morality—either reducing morality to an abstract measuring rod disconnected from practice, or emptying law of its moral dimension altogether. The aim is to challenge this dichotomy and to show that a more integrated account of moral, legal and political reasoning is both possible and necessary.

In this introduction, I will begin by offering some preliminary clarifications of the three key-terms mentioned in the title: section A focus on constructivism, while section B explores the difference between "moral" and "political" (along with "ethical" and "legal"), in order to fully understand the difference between moral and political constructivism in section C. A summary of the introduction is presented in section D followed by an overview of the entire project in section E.

[9] While I am not the first to propose this idea (Habermas, 1996) there is a limited amount of contemporary literature exploring alternative approaches. Notable exceptions include works by Liao and Etinson (2012) and Gilabert (2013), but they haven't delved into the topic as extensively as the approach I intend to take in this work, and, most importantly, they do not adopt the constructivist framework in their enquiries.

A) The Constructivist framework

When we cite constructivism in practical philosophy,[10] we are referring first and foremost to a meta-ethical framework,[11] that became popular after John Rawls' Dewey Lectures on "Kantian Constructivism in Moral Theory", wherein he offered a reinterpretation of Immanuel Kant's moral philosophy and its relevance for contemporary political debates.[12] The central idea behind moral constructivism in general is that reasons, values, and all normative facts,[13] are not discovered or revealed, but rather constructed by human agents.[14] This metaphor of construction, which carries significant implications,[15] is deployed in contrasts to the metaphor of discovery, which suggests

[10] In other branches, like philosophy of science and philosophy of mathematics, or in other fields like anthropology, sociology, pedagogy and psychology, the term constructivism is utilized in a slightly different way. Nonetheless, what all these approaches share is the adoption of the image of construction against the image of discovery. Even if some scholars have tried to locate the core of moral constructivism in is practical scope, the refutation of moral realism still remains its core, as Onora O'Neill have underlined: "Rawls contrasts constructive positions with 'intuitionism', which he characterises as impractical (so not constructive) because it prescribes a plurality of unranked principles, so offers no way of settling moral disputes. This early, thin conception of ethical construction leads Rawls to classify both Utilitarianism and his own more Kantian position(s) as constructive: both are contrasted with positions in ethics which offer no constructive criteria, so are not, or not reliably, practical. Some non-Kantian writers, for example David Gauthier, rightly view their own work as aiming at constructive criteria in this thin sense, although their approach is in no way Kantian. They are constructivists, but not Kantian constructivists […] If this were all that was meant by the term constructivist, it would be of limited interest. But most would-be ethical constructivists take a stronger view of constructivism. In particular, they reject or (at least) bracket moral realism. They assign no place or weight to distinctively moral facts or properties, whether natural or non-natural, that can be discovered or intuited and do not seek foundations for ethics in such facts (O'Neill, 2003, p. 320).

[11] For some critics, the attempt to understand normative concepts in terms of solutions to practical problems fails to adequately "move up to metaethics," since the discourse around problems and solutions remains inherently normative (Ridge 2012). Some constructivists have replied that this is misleading, maintaining that there is an intriguing continuity between normative discourse and metaethical inquiry (Bagnoli, 2024).

[12] Why Rawls preferred the term "moral constructivism" over "ethical constructivism" is a matter of discussion, since, as we will see in the next pages, he might not have a clear distinction in mind between the two, unlike the majority of constructivists. Some scholars have even questioned whether Rawls had a clear distinction not only between morality and ethics but also between morality and politics (Freyenhagen, 2011; Sangiovanni, 2008). Other authors, like Zhuoyao Li, argue that Rawls proposed (or implied) not only a distinction between morality and politics but also between ethics and morality (Li, 2016).

[13] Including all the spheres of justice we are going to see in the next pages— moral, political, ethical, and legal—along with other domains we won't address, such as the customary (e.g., one should say 'thank you' after receiving a gift) and the aesthetic sphere (e.g., this painting should be admired for its beauty).

[14] "Morality is not, as realists assume, a branch of epistemology, and its procedures are not discovery procedures, aiming to bring to light pre-existing moral facts; morality is practical, and its procedures construct moral norms which would not exist, or would not have validity, otherwise" (Rawls, 1980, p. 524).

[15] It implies at least agents who do the constructing, materials for the construction, a method or procedure for carrying it out, and a plan (O'Neill, 1989, p. 11). Still. is strongest advantage is that "unlike other metaphor used in

the existence of independent moral truths to be found: the main target of moral constructivism is thus moral realism, although it immediately presents itself as an alternative also for moral relativism[16]. Simultaneously, the construction metaphor distinguishes itself from the notion of creation *ex nihilo*, as the act of constructing is neither arbitrary nor without constraints, but simply that the objectivity of constructions "is of practical significance to human being" (Bagnoli, 2022a, p. 4).

Perhaps the clearest way to grasp these distinctions is by considering Shafer-Landau's adaptation of the question posed in Plato's Euthyphro: 'Is something good because it is desired, or is it desired because it is good?' For instance, is health considered good because we desire it, or do we desire it because it is inherently good? (Shafer-Landau, 2003, p. 42). A moral anti-realist and a moral constructivist are more inclined to endorse the first alternative: something is good because it is desired and thus moral attributes (goodness) result from a human process (desiring).[17] In contrast, a moral realist is more inclined to align with the second alternative: something is desired because it is inherently good, so moral reality precedes and exists independently of any human process. To illustrate the disparity between the constructivist and the realist, one can posit that, according to the constructivist, moral reality emerges as the outcome of a constructive function carried out by (human) beings. In contrast, the moral realist contends that there exists a moral reality independent of any specific standpoint of (human) beings—a reality that these beings endeavor to represent rather than construct. Consequently, in the event of a mass extinction of all humankind, the constructivist would argue that there would be nothing that could be deemed good or bad, while the realist would maintain that the qualities of good and bad would still persist in the world, even without any self-reflective being.

In this ontological sense, constructivism aligns with anti-realism. However, the assertion that there are no objective moral entities like "good" and "bad" existing independently of a human

---

metaethics, constructions brings to light the generative and dynamic dimension of practical reason" (Bagnoli, 2022a, p. 66).

[16] "Somewhere in the space between realist and relativist accounts of ethics there is said to be a third, distinct possibility. One such position, allegedly both anti-realist and anti-relativist, is John Rawls' 'Kantian Constructivism'". (O'Neill, 1988, p. 1)

[17] This is merely a metaphor, and most constructivists would argue that desire plays only a secondary or even tertiary role in this context, with reason being still the primary force that constructs and determines the notion of the good.

mind, does not negate the existence of objective moral facts from an epistemic standpoint. Indeed, epistemically moral facts result from a procedural construction of principles, and in Kantian constructivism this is the formalization of a moral agent's reasoning for a given practical problem. This procedure commences with non-moral and uncontroversial facts, and it is precisely from these that it constructs—hence the name—epistemically moral facts; thus, to obtain moral judgments, both the factual element (the construction materials) and the human reason reasoning on it with its constitutive principles (the procedure) are necessary[18]. The task of Kantian constructivism is to establish the authority of moral claims by reference to an account of practical reason, without relying on the agents' actual desires and preferences (O'Neill, 1989, p. 188) or their conception of the good (Bagnoli, 2022a, p. 1)[19]; the norms of reason are thus held to express a shared ideal of rational agency marked by autonomy (Reath & Timmermann, 2010, p. 4) which is constitutive and ineliminable from our practical deliberations (Korsgaard, 2008, pp. 13–15).[20]

[18] "Kantian constructivism holds that moral objectivity is to be understood in terms of a suitably constructed social point of view that all can accept. Apart from the procedure of constructing the principles of justice, there are no moral facts. Whether certain facts are to be recognized as reasons of right and justice, or how much they are to count, can be ascertained only from within the constructivist procedure, that is, from the undertakings of rational agents of construction when suitably represented as free and equal moral persons" (Rawls, 1980, p. 519).

[19] "Rawls believes that the main features of Kant's conception of practical reason can be preserved while minimizing the metaphysical implications, adopting the metaphor of construction to illustrate the very activity of reason, which is governed by a formal norm: the requirement of universality. The Kantian promise is that by appealing to such a norm about how to reason, rational agents can agree on what action to take, although they may be driven by diverse interests and concerned with contrasting conceptions of what constitutes the good life" (Bagnoli, 2022a, p. 1).

[20] The focus on the constitutive elements of reasoning is a significant recent development that has led many constructivists to identify themselves as both constitutivists and constructivists. However, it is important to clarify that it is possible to subscribe to one without necessarily embracing the other. The term "constitutivism" has sometimes even been used more broadly to include approaches traditionally skeptical of inescapable and necessary moral norms, such as those of Hume and Nietzsche. However, the more common use of the term focuses on branches of constitutivism that explicitly seek to capture the unconditional necessity of moral norms. The two most representative versions of this restricted widespread meaning are Kant and Aristotle. The neo-Kantian version holds that what essentially defines us is the concept of a rational or autonomous agent and, as a result, the fundamental practical norms are understood as constitutive standards for rational and autonomous agency. The neo-Aristotelian approach, inspired by the work of G.E.M. Anscombe and Philippa Foot, posits that the fundamental standards for living and acting well derive from the specifically human "form of life" (Anscombe, 1958; Foot, 2003). These differing views of what defines us essentially are paired with distinct conceptions of constitutive norms: the neo-Aristotelian view takes the relationship between a life-form and its exemplars as the model for constitutive norms, whereas the neo-Kantian view emphasizes that the self-constitution of an autonomous agent is sui generis, meaning it is a unique kind of process. While both neo-Kantian and neo-Aristotelian forms of constitutivism have garnered significant attention—particularly the former in meta-ethics and debates on practical rationality, and the latter in the context of virtue ethics—it is notable that the relationship between these two approaches and their differing conceptions of constitutive norms has not been explored more thoroughly in the literature (Haase & Mayr, 2019). Throughout the project, I will primarily refer to Kantian constitutivism, unless otherwise specified.

The perceived appeal of constructivism as a moral theory often revolves around its potential to provide a minimalist explanation of moral objectivity, preserving the advantages of non-naturalist realism while avoiding its associated epistemological and ontological drawbacks (Shafer-Landau, 2003)[21], but its true core lies in resolving tensions within apparent features of normative truths, addressing issues of objectivity (Bagnoli, 2015; Wong, 2008), knowledge (Bagnoli, 2018a; Engstrom, 2013), and practical significance (Bagnoli, 2021b; O'Neill, 1989), including authority (Bagnoli, 2019; Darwall, 2013c).[22] Constructivists argue that traditional metaethical theories fall short in explaining these features due to a misunderstanding of practical reason and contend that to tackle metaethical problems related to ethical judgments and normative truths appropriately, one must begin with an account of practical reason, an aspect often overlooked by standard metaethics (Bagnoli, 2013a).

B) Moral, Ethical, Legal and Political Contexts

Before delving into the differentiation between moral and political constructivism, it is useful to first clarify the general concepts of morality and politics and how constructivists interpret them. However, to approach this appropriately, we cannot restrict ourselves to the moral-political dichotomy, which in some contexts oversimplifies the rich varieties of practical reasons and the kinds of obligations we face. Instead, we must introduce a fourfold framework encompassing morality, ethics, law, and politics, which are all distinct, but connected, contexts of normativity (Forst, 2002),[23] each with its peculiar nexus of reasons and practical aims. Throughout the project,

[21] However, this is not unique to constructivism, since some versions of naturalism realism also promise objectivity without the drawbacks of nonnaturalist realism (Darwall et al., 1992). Moreover, objectivity defended on non-ontological grounds has also been a concern of antirealism for a long time (Gibbard, 1990; Hare, 1963b).

[22] Since constructivism builds moral principles through practical reasoning and shared procedures, those principles are inherently action-guiding: they are meant to guide how we live together, resolve conflicts, and justify demands on each other. This is why they have practical significance.

[23] These are not the only contexts of normativity, but they are the only practical ones—contexts that pertain specifically to practical questions. A domain like science, by contrast, falls within a theoretical context of normativity. All other practical contexts can either be included within or subsumed under these four. For example, the context of religion is, according to Forst (Forst, 2002, p. 70) subsumed under the ethical context—the context concerned with the good life, as we will explain in the next pages.

we will encounter and further clarify these distinctions, but for now, it will be helpful to have a schematic comparison of these elements, analyzing them in pairs.

Morality-Ethics

Some authors have distinguished ethics from morality by framing ethics as a wider concept that encompasses morality (MacIntyre, 1981; Rorty, 1989; Williams, 1985); in this sense, ethics pertains to a comprehensive inquiry into the fundamental question of *"How should I live?"* while morality addresses the narrower domain of moral obligation, specifically the question *"What are my duties?"*. Constructivists do the opposite by considering morality broader than ethics (Habermas 1984; Forst 2002; Bagnoli 2014);[24] morality, imbued with deontic elements on how to get along with other agents, mark the path for ethical questions on how to live a good life.[25] To answer the ethical question *"How should one live?"*, one must first engage with the moral question *"What do we owe to each other?"*. This distinction seems to result in a separation of the ethical principles (*virtues*) that regard the *good* (for me) and moral principles (*norms*) that regard the right (for all).[26]

Such separation is often framed as a contraposition between universalists and particularists (Forst, 2002, p. 155; O'Neill, 1996, p. 16).[27] Particularists argue that ethical life cannot be

[24] Saying that morality is broader does not imply that it necessarily encompasses ethics, but only that it possesses a broader normative significance or axiological priority. Yet there is also a meaningful sense in which morality can be regarded as narrower than ethics, for it concerns only the restricted class of norms that apply universally to all agents — a set far smaller than the multitude of ethical values an individual might choose to adopt.

[25] This distinction is adopted more consistently by Kantian constructivists. Bagnoli (2022), however, acknowledges that the use of these terms is, to some extent, stipulative, emphasizing that the crucial point is to employ the distinction coherently and in alignment with the intended purpose. This is why one of her books is titled *Ethical Constructivism* rather than *Moral Constructivism*, as she seeks to encompass constructivist theories such as neo-Aristotelian approaches which do not adopt the ethical/moral division exactly in these terms, but agree with scholars like Williams that "morality" carries a more restricted meaning (Williams, 1985).

[26] However, as we will see in Chapter Two, this relationship is more complex and less rigid than it may initially appear.

[27] A distinction that leaves numerous methodological problems, as underline by O'Neill: "their proponents seem content to relegate whichever domain of ethics fits less happily with their preferred view. Particularists not only reject universalist positions, but often seem content to say little about justice; universalists not only reject particularist positions, but often seem to accept that they cannot offer much of an account of the virtues. Contemporary universalist and particularists converge in the conclusion that concern for justice and for the virtues must be antithetical" (1996,

adequately captured by universal moral systems or fixed principles, but requires a responsive and context-sensitive engagement with what is morally good, recognizing that judgments about what is right depend intimately on the particular details of each situation; the connection between the good and the right, for them, is fluid and dynamic, with the right action emerging from a nuanced understanding of the good in context rather than from predetermined rules. Universalists, instead, prioritize the right over the good, emphasizing universal norms over contingent virtues.[28] What distinguishes a moral norm from other types of normative concepts, like the rule of a game, is that morality binds all persons.[29] And it binds all person precisely for its universalistic core, whether the normativity of that core is rooted in divine authority (Aquinas), rationality (Kant) or human nature (Aristotle).[30]Although critics of moral universality are diverse, their particularist criticisms share striking similarities: communitarians, virtue ethicists, postmodernists, certain feminists, Wittgensteinians, and Nietzscheans, have all attacked universalist morality on remarkably similar grounds since around 1980 (O'Neill, 2000, p. 3). Broadly speaking, these critics argue that moral thinking based on universal principles is overly abstract and insensitive to contextual differences, overlooking the particular identities of persons (Williams, 1981, p. 13), and that focusing on principles inevitably emphasizes rules and blame instead of virtues (Nussbaum, 1986, p. 4).

Constructivists avoid reconstructing the debate in this way. While it is possible to distinguish between morality as the spheres of what is right and ethics as the spheres of what is good, this does not mean that the right "phagocytize" the good. Constructivists assert that the moral

---

p. 16). Forst refers to universalists and contextualists, but he has in mind the same concerns of O'Neill (Forst, 2002, p. 155).

[28] Here we can already see why this distinction is not exhaustive. One of the most prominent moral theories, utilitarianism, is clearly universalist — at least in the sense that it holds ethical life can be captured by a single moral system, and is therefore decidedly not particularist — yet it does not prioritize the right over the good.

[29] Some particularists may accept this description, which helps distinguish morality from other cultural customs like etiquette, yet they still argue that it is entirely possible for different societies to have quite different moralities. This perspective is known as a descriptive theory of morality, as opposed to a normative theory of morality (B. Gert & Gert, 2020). A descriptive theory of morality holds that morality binds all persons within a specific society, whereas a normative theory of morality asserts that morality binds all persons universally. Constructivists subscribe to the normative interpretation. For an analysis of normative concepts in a constructivist framework, see Chapter 2.

[30] The figure of Aristotle demonstrates the complexity of the distinction between universalists and particularists, as he has been claimed by both sides (O'Neill, 1996, pp. 14–15). Aristotle's work has been interpreted and utilized by universalists to support the existence of universal moral principles (W. D. Ross, 1931) as well as by particularists who emphasize his focus on contextual and situational ethics (Wiggins, 1987b).

sphere is the context for justifying strictly universal moral norms among individuals as moral agents and members of the universal moral community (Forst, 2002, p. 237; Richardson, 2018, p. 63) but without claiming that questions regarding the good are always of particularist nature (O'Neill, 1996, p. 63). The universalist component is realized through the universalizability of our moral obligations to one another (Bagnoli, 2005, 2014), assessed through principles that are internal to and constitutive of the very exercise of rationality (Korsgaard, 1996b, 2008), like the criteria of generality and reciprocity (Forst, 2002, 2012). Ethical inquiries into how we should live, not merely to avoid harming others but for our own pursuit of happiness (the good life) can still be assessed universally, albeit by examining whether they violate the thresholds of the constitutive norms of morality and Reason.[31]

Law-Morality

If the main distinction between morality and ethics is framed in terms of what is right versus what is good, then how should we understand the context of law, which also concerns itself with what is (legally) right to do? This is why law is often grouped with morality within the broader *deontic context*, as distinct from the ethical or *axiological context*. These distinctions can be illustrated by the differing nature of the judgments they involve. Our axiological judgments—judgments about what is good or valuable—seem inevitably to operate on a sliding scale: we naturally rank actions or states of affairs in terms of comparative value, with some being better or worse than others.[32] In contrast, our deontic judgments—those concerning rightness and wrongness—do not appear to work the same way: to act wrongly is to violate something (a rule, a

[31] Henry Richardson specifically underlined how the good (what is choice-worthy) and the right (the correct way of acting) cannot be fully understood in isolation. To specify what is right, one often needs to consider what is good, and vice versa. Our understanding of moral principles involves precisely a dynamic process of specifying the good and the right in the light of each other (Richardson, 2018, pp. 32–33).

[32] While this may seem an intuitive position to hold, there are, even in this case, several hidden complications, and many would be unwilling to endorse the claim that there exists a single sliding scale of value by which all good things can be measured. For a discussion of the issue of incommensurability, see section 2.2.3.

law, or an obligation), but it would be conceptually odd to describe one wrong action as "more wrong" than another, or one permissible action as "more permissible" than another.[33]

Constructivists—and Kantians more broadly—while recognizing the analytical usefulness of distinguishing between axiological and deontic judgments, often argue that moral judgments, too, can exhibit a kind of gradation. This is evident in distinctions such as perfect and imperfect moral duties, or in the recognition of supererogatory acts, where notions of "choiceworthy" and "oughtworthy" intertwine to permit space for the notion of what is "praiseworthy".[34] By contrast, legal judgments tend to be more binary: an action is either legal or illegal, depending on whether it conforms to established legal rules. The law, by its nature, requires clear, enforceable boundaries, which explains its "black or white" character.

As such, the distinction between morality and law is usually expressed in the types of institutions they depend on: morality is a "peculiar institution" (Bagnoli, 2014) that applies to all persons, irrespective of social or cultural context; law, on the other hand, derive its authority from specific legal systems and are valid only within particular legal communities (Habermas, 1996).[35] This, however, does not imply that morality determines law, as some jusnaturalist theories affirms. From a constructivist perspective, the two contexts are indeed distinct: the law cannot be reduced to morality, nor can morality be fully absorbed into the legal domain. Yet — and this is where constructivists also distance themselves from juspositivism — the two domains are interdependent in a crucial way, because there is a core of law, that of justice, that needs to be justified morally

[33] As Keving Kinghorn aptly exemplified: "[M]urder is worse than theft in that it is "more bad" for the victim. But if wrongness is conceptually linked to obligations and to someone's rights, then either I have or I have not violated an obligation to respect another person's rights. There is no sliding scale here. Accordingly, while murder is "more bad" than theft, it does not make sense to think of it as "more wrong." (Kinghorn, 2021, p. 12)

[34] In the first chapter, specifically section 1.2.2, I explain more in detail how these categories are systematized by constructivists, recovering four categories of obligations from Kant's framework along two dimensions: universal-special and perfect-imperfect duties (O'Neill, 1989, p. 224).

[35] Habermas makes this distinction also substantive in their contents, by noting how moral norms regard general features like equality, while legal norms (or at least modern laws) are mainly about distributing individual freedoms—defining private areas where everyone is free to act as they choose. According to Habermas, this already enables us to distinguish between moral rights that emerge from the duties others owe us—duties grounded in respect for persons or moral equality—and legal rights and duties that are co-defined, arising from the need to balance and constrain liberties in a shared legal space (Habermas, 1996, pp. 92–101). In this sense, legal duties are not merely corollaries of moral obligations but are institutionalized constraints, defined by the mutual recognition of legally guaranteed freedoms.

through general principles of reasons. This is why the notion of unjust laws remains a crucial notion for constructivists.[36]

Ethics-Law

If we understand the distinction between morality and law as above, which is neither of simple derivation nor of identity, we also understand that discussions regarding the ethical neutrality of the law (namely that law should not be based on any ethical conception of the good) most of the time miss the point, because they confuse morality and ethics. Law is neutral insofar as the moral threshold is not an ethical threshold based on ethical values, but on moral norms.[37] Hence, neutrality is to be understood as a principle of *impartiality* for the justification of law: only law that does not seek a particular ethical language can properly satisfy this (moral) demand (Forst, 2002, p. 47).[38] In this sense "the legal-'negative' freedom to act and the ethical-'positive' freedom of self-realization do not exclude but imply each other: law is the protective cover for ethical identity." (Forst, 2002, p. 233).

[36] It might be argued that a theory grounded in Kantian thought faces difficulties reconciling his position on unjust laws. Kant famously maintained that morality (*Sitten*) encompasses both ethics (virtues not enforceable by the state) and law (rules that are externally enforceable). However, he appears to assume that these laws are inherently just—implying that any rule falling outside the bounds of morality does not truly qualify as law. This raises a significant problem: how can we make sense of, or critically assess, the existence of actual unjust laws if Kantian theory acknowledges only just laws as genuine law? And what about the passages where Kant says that citizens must never rebel or resist their government, even under tyranny, because justice (and rights) only exists within a lawful political system? Many responses have already been given by notable Kantian scholars (Formosa, 2008, p. 170; Ripstein, 2009, p. 209) but I will also return to the topic in greater depth in Chapter 4. For now, I wish to emphasize that the concept of "unjust law" is not only perfectly reasonable, but also analytically necessary—especially as we are, regrettably, confronted with such laws with increasing frequency today. There is, therefore, an important distinction between Kant's normative-aspirational notion of law, which identifies law with justice, and a descriptive use of the term that we must also adopt for the purposes of legal and political critique and practice.

[37] As Forst succinctly put it: "principles of law constitute not a particular form of life but a system of justified norms that, to be sure, is not without ethical demands (and not entirely without ethical content) but is not justified on ethical grounds" (Forst, 2002, p. 67)

[38] "Since the neutrality principle relates primarily to the validity criterion of general norms, it does not imply neutrality in the process of justification in the sense that ethical arguments would be excluded from it […] what is important is that ethical arguments, when they propose values as the foundation for general regulations, be "translatable" into universal arguments. They must be compatible with the principle of public justification; however, they do not therefore completely detach themselves from the ethical background from which they originate. The criteria of strict and restricted generality do not disengage arguments or justified norms from their social contexts." (Forst, 2002, p. 47).

Laws are, therefore, not the arbitrary legal actualization of specific ethical values, but they are the outcome of deliberative justificatory processes that must respect the moral threshold based on constitutive principle of reasons, which limits the possibility that specific ethical views become the foundation of general law. At the same time, law must protect the pluralism of ethical life: since there are many different and often incompatible conceptions of what constitutes a good or fulfilling life, morality cannot declare a particular form of life "false" simply because it cannot be universally realized. This is why it for constructivists a central notion is that of "reasonable ethical pluralism" (Bagnoli, 2018b).[39]

In practical terms, this has significant implications for how a just state should approach deeply contested ethical matters. For example, just state has to find a justification on ruling on abortion with just rights: this implies ruling without appealing solely to a particular ethical doctrine (e.g., a religious accounts of personhood), but formulating laws that protect the fundamental rights of all affected parties — including the pregnant person — in a manner that does not impose one ethical conception of life over another. The goal is not ethical consensus, but moral legitimacy grounded in justifiability to all.

Law-Politics

If Law, as a "protective cover", reflects the fragility of ethical identities, law claims to generality has to be legitimized not only institutionally, but also within democratic political communities, through the discursive deliberation of citizens who are, at the same time, both authors and addressees of the law. Citizens are not merely passive recipients of laws, they actively contribute to their formation as authors; as such, they bear responsibility for shaping the legal and political structures of their society which is different from their legal responsibility. This means that laws are appropriate only when they conform to a constitution that all citizens can reasonably accept.

[39] In contrast, the idea of a "reasonable moral pluralism" is rejected, since morality, by constructivist definition, demands universally justifiable principles.

Legal autonomy (as addressees of the law) and political autonomy (as authors of the law) are considered "equiprimordial"—that is, they are "co-originally" fundamental (Habermas, 1996, pp. 104, 127).[40] Both are equally presupposed by the legal institutionalization of a democratic political order, since the freedom of a subject in a constitutional state is the freedom to "obey no law other than one to which they have given their own rational consent." One is not truly free unless one is both subject to and author of the law at the same time (Forst, 2024, p. 204).[41]

Politics-Ethics

The relationship between ethics and politics has long been at the heart of the communitarian critique of liberal and pluralist conceptions of democracy. Communitarian theorists have argued that a political community cannot be sustained solely on the basis of procedural norms or legal rights; rather, it must be grounded in a shared ethical conception of the good—a thick sense of collective identity, purpose, and belonging.[42] The core of this critique lies in the concern for the "sociocultural presuppositions of an intact democratic community" in which the question of necessary institutions takes a back seat to the focus on the ethos of democracy—

[40] Forst distinguishes himself from Habermas, who identifies two sources of this equiprimordiality: the principle of discourse and the form of law. For Forst, by contrast, there is a single source of both forms of autonomy: the principle of justification (Forst, 2012). The relationship between moral autonomy and legal/political autonomy is immanent; the former serves as the logical core of the latter, without being fully absorbed into it.

[41] "Imagine a perfect freedom-regulating computing machine that could properly process all justifiable freedom claims and calculate a right, generally justifiable measure that only needed to be secured by law afterwards. Would that be a Kantian regime of freedom? No, it would not, for in such a regime freedom would have come about in a heteronomous way. Legal freedom cannot be thought of or exist without the exercise of political liberty in the form of justificatory autonomy as a co-legislator. The freedom of a subject in the legal state is the freedom "of obeying no other law than that to which he has given his consent." This is the real meaning of legal independence or political–legal non-domination: you are only independent or not dominated if you are at the same time subject and author of the law, for otherwise your right to freedom or to justification is only half realized. You might receive freedom or some justification, but if you cannot be the co-author of the laws of freedom, it is not real freedom that you are being offered, only a dependent or possibly dominated, and in any case heteronomous, scope for action." (Forst, 2024, p. 204).

[42] For communitarians such as Michael Sandel, the political community is understood as a constitutive community—a form of association that is not merely contractual or procedural, but one that is prepolitically integrated through shared historical, cultural, and ethical commonalities. (Sandel, 1982). In this view, political identity is a thick, ethical identity, and political obligations are not freely chosen but are rather embedded in one's social identity and moral development. The risk of not recognizing this central feature leads, for Sandel, "toward individualism, privatization, and ultimately, the rise of despotic or undemocratic rule over a mass of fragmented legal subjects (Forst, 2012, p. 168). For a rebuttal of this position see Forst (Forst, 2002, pp. 9–15, 2012, pp. 165–174)

that is, "the manner in which citizens conceive of themselves as members of a political community, what they have in common and what responsibilities they bear" (Forst, 2002, p. 88).

From a constructivist perspective, the requirement of a shared ethical conception of the good as the basis for political legitimacy is both empirically problematic and normatively exclusionary. It fails to account for the reasonable pluralism of ethical worldviews in modern democratic societies and risks marginalizing minorities who do not share the dominant conception of the good. Instead of grounding political legitimacy in shared ethical substance, constructivists argue that it must rest on a procedural principle of justification that all citizens, regardless of their ethical commitments, can reasonably endorse: political discourse is not (primarily) an ethical discourse of collective self-understanding, but rather a process of justification aimed at identifying universally acceptable norms, while still recognizing and respecting diverse ethical identities. Similarly, democratic legitimacy is not based on ethical unity but on the reciprocal justification of laws and institutions among free and equal persons; the goal is not to dissolve ethical particularity into a shared ethos, but to enable coexistence within a framework of justice that protects autonomy—both legal and political—through the equal distribution of justificatory authority.

If reasonable disagreement is an inevitable feature of a free society — as many scholars, including non-constructivists (Gray, 2000; Gaus, 2010; Larmore, 2020),[43] acknowledge now—then the central task of political philosophy is precisely to address the pluralism of ethical values. This challenge gives rise to distinctively political problems: determining the conditions under which authoritative and enforceable rules for managing such conflicts can be justifiably established and legitimately recognized. In short, it is the problem of power (how it is exercised, justified, and constrained).

Politics-morality

We are now equipped with the conceptual tools necessary to address the central relationship explored in this project: the relationship between morality and politics. While these

[43] "[A] leading theme in modern thought is that reasonable people, exercising their general capacities of reason on ethical questions as best they can and in good faith, tend to come to different conclusions—not through any defect of reason, but instead because of their different backgrounds, their different senses of what is salient, and their different ways of weighing disparate considerations" (Larmore, 2020, p. 36)

two “contexts are partially (not completely) congruent” (Forst, 2012, p. 111), they operate according to distinct procedures and serve different functions within the framework of a just society. While morality gains legitimacy through its adherence to universal principles and its capacity to be justified through general reasonable deliberation, political decisions and policies are justified within a specific community through democratic processes, institutional procedures, and often through negotiation and compromise among various stakeholders. Nonetheless, political legitimacy is derived not only from the ability of these processes to represent the interests of the public, but also and especially from the fairness of these processes which refers also (but not only) to moral general principles.

It is essential, however, to reaffirm that this moral grounding does not entail a phagocytosis of the political by the moral. Morality provides critical but limited guidance: moral principles determine the conditions of legitimacy—they specify the parameters within which legal and political institutions must operate—but they do not resolve every question of institutional design, legal interpretation, or policy implementation. These questions remain the domain of political reasoning, informed by context-sensitive judgment and specific empirical data. To put it another way, morality can outline the architecture of a legitimate political order, but the interior structures (the policies) are constructed politically. Morality sets formal constraints on what can count as a just political order, but it does not dictate the total content of political arrangements.

The normative framework of politics is thus both morally constrained and politically constructed—a tension that constructivists seek to resolve in different ways (Bagnoli, 2006; Forst, 2002; Habermas, 1996; Korsgaard, 1996a; O’Neill, 1996; Rawls, 1971; Richardson, 2018).[44] Some, like Rainer Forst, highlight the aforementioned formal congruence by appealing to a unified normative ground, the principle of justification, which, articulated through the criteria of reciprocity and generality, constitute the formal core of legitimate moral and political justifications (Forst, 2012, p. 6).[45] In this way, the principle of justification functions as the primary normative

[44] Kant equation of acting in accordance with the moral law and acting in accordance with the idea of a self-rule, and the interconnection of public and private right, could already be seen as a tentative attempt to close the gap between these two realms (4:431, 6:230, 8:349).

[45] “Reciprocity means that no one may refuse the particular demands of others that one raises for oneself (reciprocity of content), and that no one may simply assume that others have the same values and interests as oneself or make recourse to "higher truths" that are not shared (reciprocity of reasons). Generality means that reasons for

foundation that guides both moral reasoning and democratic deliberation, through their respective, but unified, procedures of constructions (Forst, 2012, p. 218).[46]

However, some constructivists, like Henry Richardson, seek to reject the "seductive analogy between morality and politics" (Richardson, 2018, p. 179). Richardson believes that Kant's attempt to encapsulate the entirety of morality within the framework of autonomous principles becomes untenable if it fails to appropriately account for the significant influence that other agents have in shaping the principles of practical agency. He argues that morality should not be understood as a sphere of collective self-rule, bridging the gap between morality and politics; rather, it ought to remain distinct from the political domain in order to preserve the unique kind of normativity that characterizes each. Moral authority cannot be established through the centralized mechanisms typical of democratic processes, such as the election of representatives, majority rule, and contestatory mechanisms (Richardson, 2018, p. 21), nor can it be enforced through coercive sanctions (Richardson, 2018, p. 7). [47]

Still, it seems that some local democratic set-ups have the capacity to authorize new moral norms and that, contrary to Richardson's view, a vote can indeed be one of those "legitimate social processes" that ratifies a moral norm.[48] Democratic political frameworks serve as deliberative arenas not only for discussing rights and duties but also for their formal ratification and realization. Thus, the interplay between morality and politics is more nuanced and interconnected than might

---

generally valid basic norms must be sharable by all those affected. The criterial strength of these requirements is a substantive implication of the theory I am proposing." (Forst, 2012, p. 6).

[46] "The idea that there are two separate procedures for [moral and political] construction is thus itself an abstraction: moral justification is – in a normative-formal sense – the core of political justification" (Forst, 2012, p. 218).

[47] "Although there may be moral grounds for concluding that we ought to rule ourselves politically – in territorially limited democracies – there is no imperative that morality as a whole serve us as a tool of self-rule. Indeed – setting aside Kant's distinctive equation of acting in accordance with the idea of self-rule or autonomy – there are at least three reasons why it would be a mistake to cast morality as such as a marking out a sphere of collective self-rule" (Richardson, 2018, p. 179). The three reasons that Richardson provides to avoid considering morality as a sphere of collective self-rule are that morality, first, is not upheld through coercive sanctions; second, it does not involve the formal expression of will typical of democratic procedures; and third, it is not organized through contestatory mechanisms, but rather governed by objective, universal principles.

[48] Consider recent instances in which elections in specific democracies have centered around moral issues such as same-sex marriage. When we say that an election serves as a referendum on a moral issue, we do not mean that it is literally a referendum. Instead, we mean that the electorate, as a moral community, views the contested issue as significant, closely listens to what the various parties have to say about it, and casts their votes with a sense of moral responsibility.

be initially apparent, and this is why it is crucial to properly clarify these normative contexts in order to understand the progression from moral to political constructivism. [49]

C) Moral and Political Constructivism

As we said, over the past decades moral constructivism has risen as a prominent moral theory, with philosophers crafting diverse versions inspired by historical thinkers[50]. Concurrently, political constructivism has garnered recognition as a method for formulating and defending principles of justice and legitimacy, once again thanks to the contributions of John Rawls.[51] The late Rawls emphasizes that political constructivism, representing his mature interpretation of justice as fairness, is a practical doctrine, not a metaphysical one: political constructivism allow

[49] It might be argued that what I just did in this section was not clarifying concepts but rather assigning new definitions to certain terms. However, this critique overlooks one of the crucial ways in which we can clarify a concept: by providing more precise conceptions. To highlight this nuance, we can turn to John Rawls's distinction between a concept and a conception. A concept is a broad term whose specifics are elaborated by a particular conception. For example, the concept of justice includes the idea of a proper balance, yet it is left to each conception of justice to interpret what constitutes a proper balance: "men can agree to this description of just institutions since the notions of an arbitrary distinction and of a proper balance, which are included in the concept of justice, are left open for each to interpret according to the principles of justice that he accepts" (Rawls, 1999, p. 5). My aim was thus precisely to provide some staple conceptions for terms like morality and politics by referring to the constructivist framework, in order to develop a constructivist conception of human rights.

[50] For example, along with the Kantian varieties, one can now find Aristotelian (LeBar, 2008), Humean (Lenman, 2010; Street, 2008), Hegelian (Laitinen, 2008; Yeomans, 2012) and Nietzschean (Silk, 2014) versions of constructivism represented in the philosophical literature. Although I cannot delve into all the specific currents by providing a systematic literature review (as that would be a distinct project on its own), I will, throughout the text, refer to some singular authors when they offer ideas that are pertinent to the human rights question (and not just the metaethical). The central assumption shared by these authors, the reason why they define themselves as constructivists, remains the same as discussed earlier: moral, political and ethical principles are constructed, not discovered, and the process of construction is an activity undertaken by rational agents, both individually and socially. For a thorough examination of different constructivist currents in metaethics, refer to Bagnoli (2013a, 2022a).

[51] Today, however, the term 'political constructivism' is even more popular in the field of international relations, where it takes on a slightly different meaning from its philosophical roots, by incorporating influences from anthropological and sociological constructivism. Nevertheless, Nicholas Onuf, who is credited with coining the term 'constructivism' in relation to international relations theory, emphasized the importance of Rawls and Habermas (and their readings of Kant as a constructivist) in understanding how political concepts are shaped by ongoing processes of social practice and interaction (Onuf, 2013). In this project, I will frequently reference the influential work of both Habermas and Rawls (section 5.2.1), but their roles will primarily serve as a conceptual backdrop against which to situate and highlight more recent scholars working within the contemporary political constructivist framework. It is important to note, moreover, that not all political constructivists—at least in the sense that concerns this project—are fully aligned with Rawls (as we will see in the next pages) or Habermas's approach. For instance, Onora O'Neill critiques Habermas's perspective for the potential risk of becoming "particularist" as it may rely on specific agreements such as actual consensus, or for requiring "metaphysical backing," given that it appeals to agreements "achieved under supposedly ideal conditions, whose status needs vindicating" (O'Neill, 1996, p. 52, note 26).

for reasonable political disagreements, by intentionally remaining neutral on deeper philosophical commitments (Rawls, 1985). Rawls believes that this shift resolves certain criticisms of his earlier formulations, particularly those raised by communitarian critics. He argues that a comprehensive constructivist account of morality would encounter reasonable disagreement, making it unsuitable as a stable foundation for enduring consensus on justice requirements. In contrast, a constructivist political theory avoids this problem, providing a more stable basis for consensus on justice.

However, despite his commitment to exclude comprehensive moral doctrines, some authors underline that he makes it clear that the political conception of justice is nonetheless a moral conception, underlining a continuation and not a separation between the two (Li, 2016). Other authors note that Rawls' transition from *A Theory of Justice* to *Political Liberalism* weakened his constructive proposal, precisely because he feared that ascribing universality to morality would be too metaphysically demanding. In contrast, O'Neill and Forst recover a recursive and discursive conception of practical reason that eschews transcendental moral truths without relinquishing moral validity, understood as a reasonable claim "that cannot be rejected with general and reciprocal reasons" (Forst, 2002, p. 238).[52]

For O'Neill, the principle of moral justification does not rely on any demanding metaphysical presuppositions. Instead, it operates on the assumption that there is no metaphysical authority over reasons;[53] the validity of its principles must be demonstrated solely in a procedural sense, as principles that are shareable among a plurality of persons (O'Neill 1989, 212-220). [54]

[52] However, Forst, while acknowledging this complementarity, also seeks to elucidate the differences between O'Neill's conception of intersubjective practical reason and Apel and Habermas's conception of vindicating communicative reason, on which he relies (Forst, 2002, p. 192).

[53] "Several features distance the alternative constructivism sketched here from Rawls's pioneering project. First, idealized accounts of agents, their rationality and their mutual independence are explicitly rejected: There is no appeal, however oblique, to transcendent moral claims. Second, abstraction from the determinate desires of agents is complete: Hence no special ingenuity is needed to avoid either heteronomy or cruder forms of relativism. Third, and as a consequence, it is impossible to answer questions about the hypothetical choices of abstract agents. Fourth, the construction therefore has to fall back on modal questions about the possible choices of abstract agents, and construct an answer to the question "What principles must a plurality of abstractly characterized agents reject?" Finally, rejection of nonuniversalizable principles can guide action by requiring that we ensure that the agents actually affected, with their particular identities and vulnerabilities, can genuinely choose or refuse those principles. If this sketch can be filled out, there is at least some space between realism and relativism." (O'Neill, 1989, p. 218).

[54] O'Neil asks: "What principles can a plurality of agents of minimal rationality and indeterminate capacities for mutual independence live by?" The answer she arrives at is that: "No plurality can choose to live by principles that aim to destroy, undercut or erode the agency (of whatever determinate shape) of some of its members. Those who become victims of action on such principles not merely do not act on their oppressor's principles; they cannot do so.

Moral principles of justice are not justified solely by the hypothetical consensus of rational parties; rather, they emerge from what she calls the recursive structure of practical reason. This means that reason must ask of each principle whether it could be consistently applied to everyone and whether it could be reasonably accepted by others. Practical reason, in this view, test itself in a self-reflective way: this defines what O'Neill calls *moral constructivism* (O'Neill, 1989, pp. 43–44).[55] Then, the possible consent of actual agents is sought to maintain the correct balance between abstract principles and their concrete justification (O'Neill, 1989, p. 217) which forms the core of a version of *political constructivism* that acknowledges the finiteness of human beings.

In the same vein, Forst, while commending the innovative aspects of Rawls's early work,[56] aims to steer clear of certain late deviations, like the substantive presuppositions with his conception of the person, which is based precisely in failing to clearly distinguish between ethics and morality. According to Forst, Rawls at times uses the concept of the moral in an ethical sense, and at other times in a moral-reasonable sense. This ambiguity obscures the fact that a conception of justice can be both morally justified and ethically incorporable while at the same time maintaining a distinction between the two spheres.[57] Rawls's use of the concept of political reason "does not underscore clearly enough that the overlapping consensus is a moral consensus to which

---

Victims cannot share the principles on which others destroy or limit their very capacities to act on principles." (O'Neill, 1989, p. 213).

[55] "If reasoning has only a discursive and recursive grounding, and lacks transcendent vindication, then even the "supreme principle of practical reason", the Categorical Imperative, has no greater authority than that it is a principle capable of guiding the interactions, including the communicating, of beings whose coordination is not naturally guaranteed. The Categorical Imperative states essential requirements for a possible community (not an actual community) of separate, free and rational beings." (O'Neill, 1989, pp. 43–44)

[56] "In my approach, constructivism means that the principle of justification, which specifies, depending on context, which norms are to be constructively justified in which way, must not only be normatively binding (as in Kant, and in discourse ethics, albeit in an ambivalent sense), but also that it cannot itself be constructed, but has a reconstructive character. The underlying normativity of this principle grounds the normativity of the (justified) norms constructed on this basis. Rawls expressed this with exemplary clarity in his pioneering interpretation of Kantian constructivism. According to Rawls, Kant's categorical imperative is not constructed but "laid out" as a procedure based on a reflection of practical reason as it appears in moral deliberation and judgment" (Forst, 2024, p. 7).

[57] "As a Kantian theory: Rawl's theory rests on the principle of general justification: his conception of practical reason is explicated on a fundamental level on the basis of the "ideas" of reason, that is, the moral person and social cooperation. These serve the justification both of moral principles of equal rights and of substantive principles of social justice (for which certain additional assumption are necessary); and in their twofold task lies the reason for the political-moral double character of Rawl's model" (Forst, 2002, p. 238).

no ethical claim to validity as such is connected and whose 'reasonable' validity is therefore not dependent upon being transformable into an ethical 'truth'." (Forst, 2002, p. 185).[58]

Political constructivism in Rawls consists thus of three steps. 1) The reflective reconstructions of the principles and ideas of practical reason (conceptions of reason, of the person and of society).[59] 2) Their employment in laying out a procedure which embodies these principles (original position). 3) The constructions of principles of justice with this procedure (well-ordered society). In so doing "the principle of justice cannot be reasonably rejected by "reasonable" persons" (Rawls, 1993, p. 124). Forst argues that structuring political constructivism in this manner is both stronger and weaker than a purely procedural theory. It is stronger because it makes substantive assumptions about conceptions of the person (such as primary goods), but it is weaker because these context-dependent assumptions cannot claim to be as "reasonably" justified as the fundamental conceptions of practical reason.

Forst instead consider political constructivism as complementary with moral constructivism and delineate the three levels of normativity which characterize the steps from moral constructivism to political constructivism. (Forst, 2024, pp. 10–11). 1) At the basic level, there is the binding power of the constitutive principle of justification, which establishes both the duty and the right to reciprocal-general justification in morally relevant practical contexts. 2) At the second normative level there are moral norms generated by strictly reciprocal-general justification, involving a procedure of moral constructivism. This process combines forms of factual and counterfactual justification, fostering discursivity among those affected and reflecting on justifiability within a comprehensive moral community of all human beings. 3) Finally, there is the political-legal level of political constructivism. Here, human rights and basic principles of justice must be grounded to form a basic structure of justification wherein individuals subject to a normative order become part of the governing authority while being protected as legally,

[58] This distinction becomes clearer when we acknowledge that the primacy of justice over ethical doctrines cannot be upheld with ethical arguments: "just action is demanded not for the sake of the good of a particular ethical doctrine but for the sake of equal respect for the legitimate and 'reasonable' claims of all. Justice might be a diamond that reflects different colors in the light of different comprehensive doctrines; its moral worth is not however located in this ethical reflection" (Forst, 2002, p. 185).

[59] "The principles of practical reason are therefore to be found through reflection on what it means to think and judge rationally and reasonably, and the ideas of practical reason, namely, the conception of person and society, are to be found through reflection on the subject and practical context of this thinking and judging" (Forst, 2002, p. 187).

politically, and socially non-dominated equals through basic rights. Beyond basic rights and principles, political-legal norms stemming from this basic structure of justification require further justification, especially when originating from majority opinions, with those whose non-rejectable claims have been overruled retaining a discursive veto right. The aim is to achieve a justified basic structure, progressing towards the ideal of a *noumenal republic* where the subjected are both governed and law-givers, constituting political and moral progress (Forst, 2024, p. 13).

To summarize, the conceptions of the legal, political, ethical and moral person must be formulated "in such a way that the abstraction that necessarily accompanies these conceptions always relates to the possibility of protecting and recognizing concrete identities and reflects the conditions of realizing the just in communities" (Forst, 2002, p. 239) through a conception of practical rationality suitable for finite rational beings (O'Neill, 1989, p. 74). Practical reason is precisely the faculty of providing reasons which are appropriate to the specific normative contexts.[60] Moral and political constructivism are the steps towards this idea of a justified and just society. While moral constructivism develops a conception of rights as universal demands, political constructivism develops conceptions of legal and social structures in which these reasonable claims are institutionalized and realized as basic rights.

D) Human Rights at the Intersection of Justice Contexts

The reason why I have dedicated significant effort to clarifying these concepts, delving into all the spheres of ethics, morality, law and politics, is due to their intrinsic connection with the idea of human rights. The complexity of practical judgment in different normative domains highlights the need for intersubjective justification, where actors justify their actions to others based on ethical, legal, political, or moral grounds. And as domains of activity, "the reasoning that we bring to them must be practical reasoning, that is reasoning which we and others can use both in personal and in public life not merely to judge and appraise what is going on, not merely to

[60] "A theory of justice must not absolutize one of these dimensions and shape the others according to it: justice maintains the boundaries between these spheres by giving effect – in accordance with generally justified norms – to ethical identity, equal rights, political membership, and moral respect. […] The person at the center of a theory of justice is not just an ethical person, a legal person, citizen, or moral person, he or she is all of these simultaneously in a different way: he or she is ethically, legally, politically, morally autonomous" (Forst, 2002, p. 240).

assess what has been done, but to guide activity" (O'Neill, 1996, p. 2) The integration of these different spheres of normativity is essential for responsible and reasonable behavior, both individually and within the broader societal structure. This integration reflects a comprehensive sense of justice that respects ethical autonomy, personal and political liberty, and moral recognition of people within their communities.

Consider the example of famine: the occurrence of hunger-related deaths in underdeveloped countries is not merely a factual event but is imbued with deep moral implications that call for practical deliberation to address systemic injustices on a global scale.[61] In the current global landscape, international politics has led to the development of legal instruments like human rights, specifically designed to address such issues. Because access to adequate food is recognized as a human right, governments and international bodies have a legal duty to prevent famine through policy, resource distribution, and emergency response. Where famine occurs, it often reflects a failure to fulfill these legal obligations—whether due to negligence, corruption, or war; human rights provide individuals and organizations with both a language and the means to challenge the systems that cause or tolerate famine, demanding remediation, transformation, or overcoming of these injustices.

Therefore, famine, when analyzed through the lens of human rights, engage all these four "contexts of justice" (Forst, 2002, 2010). [62] Human Rights simultaneously safeguards the practices

[61] "Our moral intuitions of the wrong of extreme global poverty and inequality derive from the prior notion of justice and injustice as relational concepts. However, social relations are intersubjective relations, therefore the objective wrong of inequitable social conditions is exacerbated by the fact that it is the result of someone's action or inaction. The intersubjective component of social injustice makes not just a bad situation, but also a moral problem. Therefore, while it is correct to consider social injustice as objectively bad, it is also important to bear in mind that such situations are the result of the expression of autonomous agency by some parties that leads to such adverse results for others." (Nwaneri, 2017, pp. 222–223).

[62] "Human Rights are a complex phenomenon, comprising an array of different aspects. They have a moral life, expressing urgent human concerns and claims that must not be violated or ignored, anywhere on the globe; they also have a legal life, being enshrined in national constitutions and in lists of basic rights, as well as in international declarations, covenants, and treaties; and they have a political life, expressing standards of basic political legitimacy" (Forst, 2010, p. 711). Forst here is leaving aside the 'ethical life' of Human rights because he does not think that an Ethical constructivism regarding human rights is a productive path, and that ethical judgments, resting on a notion of the good life, cannot properly ground human rights. Nevertheless, he acknowledges the inherent ethical dimension within the concept. As he later clarifies, ethical justifications are not "irrelevant for the justification of rights claims; rather, they are insufficient to establish the normative character of such claims as reciprocally and generally binding" (Forst, 2010, p. 726). In the example of famine, it is clear how some ethical considerations can come into play: rigidly hierarchical societies (such as under certain historical interpretations of the caste system in India), see some persons as inherently inferior or destined for hardship, and their suffering, including starvation, is viewed as natural or ethically

for a good life that diverse human cultures worldwide have (ethical sphere) as long as they do not violate a moral threshold that is constitutive of any reasonable person (moral sphere); this is why we demand judicial tools to enforce them (legal sphere) and seek accountability from institutions responsible for their implementation in cases of failure (political sphere). [63]

E) Laying Out the Blueprint: Project Overview

This project aims to explore and elucidate how recent constructivist approaches in moral and political philosophy can meaningfully contribute to the ongoing discourse surrounding human rights. The study is structured in two parts. In Part I, I delve into the definitional challenges surrounding human rights. After exploring how different philosophical and legal perspectives contribute to a lack of consensus on the general nature, form and function of right, by analyzing different constructivist proposals (chapter 1), I provide a definition of right that can be accepted by the majority of them as: legitimate specificatory demands. After distinguishing and differentiating rights from other normative concepts, such as reasons and values (chapter 2), I argue that what makes the demand legitimate (and therefore the right normative) for a

---

acceptable. It is similarly clear how these ethical practices violate general principles of morality: famine is not a just consequence of one's birth status—it is a failure of distribution, governance, and moral responsibility.

[63] To reinforce this sketch, consider another example. When we assert that the practice of female circumcision performed on Somali girls is immoral, we are not merely expressing an ethical preference. Rather, we are making a moral judgment grounded in the recognition of the profound physical and psychological suffering this practice inflicts on young girls. This suffering is not morally defensible—it cannot be justified through reasons to the people that are victims of the practice. Our judgment does not stem solely from our belonging to a different cultural background; instead, it is based on an appeal to what appears to be a universal moral principle—namely, that it is wrong to inflict unnecessary and severe harm on vulnerable individuals. This principle, we argue, transcends cultural boundaries and should be recognized as valid by all human beings, regardless of their cultural, societal and ethical context. This is the sense in which a human right is a moral right; when we are convinced of its universality, we do not want simply to remain an abstract ideal, but a norm with legal weight, and this is why we expect governments and international bodies to codify the right into enforceable laws. In the case of infibulation, this means legally prohibiting the practice, prosecuting those who carry it out, and protecting potential victims through judicial mechanisms. If institutions fail to prevent or punish infibulation, human rights enable us to hold them politically accountable, and civil society, media, international organizations, and democratic mechanisms can pressure governments to fulfill their obligations—whether by supporting victims, reforming laws and providing education. While this may appear to many as a call for external intervention by third parties—which I reaffirm can indeed be necessary at times, and states with the capacity to assist have a responsibility to do so—I understand that it raises legitimate concerns about historically insensitive and imperialistic interventions by Western powers. However, the horrors of colonialism should not lead us to abandon people in need by relying uncritically on notions of cultural difference or non-interference. For further discussion on cultural relativism, see footnote 3 and the next one.

constructivist stems from autonomous and public reasoning about why it should be so. A claim becomes a legitimate demand only if it can be reasonably justified to those responsible for fulfilling its corresponding duties.

However, does the fact that humans are the only known beings capable of reasoning necessarily mean that they alone are entitled to certain rights? In order not to exclude animals and future AIs or possible sentient aliens, should we perhaps speak of person rights instead of human rights? This suggests the need for a constructivist conception of humanity or personhood. To this end, I explore whether the concept of dignity, often considered the bedrock of human rights, is compatible with constructivism or should be discarded (chapter 3). I conclude that it should be retained but conceptualized as a status-concept rather than a value-concept, attained through reason (as it can always be justified) rather than being self-evident. It is feasible to uphold the foundation of human rights in the concept of dignity without succumbing to foundationalist metaphysical tendencies. By elucidating how this concept has evolved through the ages and intertwining it with that of respect, I redefine it from a constructivist standpoint to include the notion that we consider every moral person worthy of respect. Yet, the question remains: can non-human animals and non-human agents be considered moral persons?

In Part II, I engage with the core question of the contemporary human rights debate: the nature of the universality to which they aspire. I critically examine the historical discourse between jusnaturalism and juspositivism (chapter 4), arguing that both rely on an erroneous, strict division between facts and values that leads to form of foundationalism or formalism both undesirable. Using constructivism as a framework, I propose to shift the central question from whether human rights are natural to whether they can possess universal validity, in a qualified sense (chapter 5). I will show that most recent trends in constructivism adopt a form of universality which seems to align with jusnaturalist tradition due to their strong commitment to avoiding arbitrariness, while simultaneously paralleling the procedural development of normativity characteristic of the juspositivist approach. I will therefore argue that the constructivist conception of human rights constitutes an original alternative that preserves the strengths of both traditions while avoiding their respective weaknesses: human rights are both institutional (positive) and universal (natural), meaning they are established by human institutions but are not arbitrary or ethnocentric: they are

grounded in the faculty of Reason, which is shared by all human beings across time and cultures.[64] This approach explains how a law can be unjust without appealing to foundationalist realism, as these principles are not independent of human reason but rooted in it. At the same time, it respects diverse cultural practices without succumbing to relativism, recognizing that beneath these practices are shared principles of reason that serve as a cross-cultural and cross-generational framework.

This makes constructivism particularly suited for addressing human rights in a globalized world. In an international context, where diverse legal, moral, and cultural systems must coexist, a rigid appeal to natural law may not resonate universally. Constructivism, by contrast, frames human rights as a practical project in which all individuals participate in reasoned deliberation, reflecting the complexity of global moral landscapes while maintaining the universality of certain norms. By grounding universality in human reason rather than immutable truths, this approach addresses classic relativist critiques of the globalization of human rights. Laws or practices that undermine human dignity or rational autonomy—such as those enforcing discrimination or subjugation—can be critiqued not for violating objective truths but for failing to align with the constitutive principles of reason that underpin fairness and justice across societies.

This project could attract criticism due to its extensive theoretical elaboration, introducing the potential hazard of overly complexifying the 'intuitive' concept of human rights and, consequently, diminishing its practical advantages. In order to deflect this critique, it is important to underline Amartya Sen's cautionary reminder: "The invoking of human rights tends to come mostly from those who are concerned with changing the world rather than interpreting it (to use a classic distinction made famous, oddly enough, by that overarching theorist, Karl Marx). It is not hard to understand their unwillingness to spend time trying to provide conceptual justification,

[64] In affirming this, I am not claiming that all cultures—past or present—have consistently and continuously embodied reason. Historical contingencies, power structures, and entrenched ideologies play a significant role in shaping how traditions are formed, sustained, and transmitted, for example through habits (Bourdieu, 1990; Dewey, 1925). However, what I am asserting is that even within such contexts, individuals retain a capacity to respond to the demands of reason. To deny this would be to deny the agency of those who, despite being deeply embedded in oppressive practices, have nevertheless resisted them. This includes, to reconnect to the previous footnote, documented cases of mothers who sought to protect their daughters from infibulation, defying the norms in which they themselves were raised. A constructivist theory of human rights needs to resist both cultural relativism and moral absolutism, by affirming that reason can speak across traditions, without erasing the historical and social complexity of how moral insight emerges. The possibility of critique and progress depends not on cultures being uniformly rational, but on the justifications that are given in favor of the practices.

given the great urgency to respond to terrible deprivations around the world. This proactive stance has had its practical rewards, since it has allowed immediate use of the colossal appeal of the idea of human rights to confront intense oppression or great misery, without having to wait for the theoretical air to clear. However, the conceptual doubts must also be satisfactorily addressed, if the idea of human rights is to command reasoned loyalty and to establish a secure intellectual standing. It is critically important to see the relationship between the force and appeal of human rights, on the one hand, and their reasoned justification and scrutinized use, on the other" (Sen, 2004, p. 317).

Appealing to human rights as an intuitively understood concept can indeed be beneficial, yet it is still necessary to delve into the philosophical intricacies underpinning these rights for them to wield true persuasiveness. Developing a well-defined theory of human rights is not merely an academic exercise, it has practical implications. The international application of human rights is fraught with ambiguities: questions such as how to precisely define "human rights", why certain values are recognized as human rights while others are not, and the nature of responsibilities associated with these rights, along with identifying the responsible parties, are all areas of ongoing debate and confusion not only at a theoretical level but also in the practical implementation of human rights. This issue was recognized not only by moral philosophers like Amartya Sen but also by scholars who adopt a more practical stance, such as Charles Beitz, arguing that such a theory should illuminate the challenges posed by international practices that give rise to these dilemmas. Both aspects, moral and political, are deeply intertwined in the discourse of human rights, and any comprehensive theory must address this intersection to provide a robust understanding and guide effective implementation (Beitz, 2004, 2009). The aim of this project is to provide a new reading of the debate through a new framework, in order to obtain a clearer picture of the issue, potentially clearer than even Bobbio envisioned.

PART I

CONCEPTUAL ANALYSIS: TOWARD A CONSTRUCTIVIST DEFINITION OF HUMAN RIGHTS

There are at least two tasks that a book devoted to philosophical aspects of human rights ought to undertake. One is the analysis of the concepts of a right and of a human right. This involves explaining the current conception of human rights, sketching its historical development, offering an analysis of rights generally, and explaining the relation of human rights to other kinds of rights. The second task involves examining the normative foundations, if any, of human rights. Here the goal is to give a plausible account of the grounds for believing that there are basic rights shared by all people.

(Nickel, 1979, p. 1233)

What is the point of trying to clarify the meaning of a term so central to contemporary global political life if one's definition is in no way tied to the way the term is actually used? Clearly a philosophical conception of human rights must be bound in some ways to the way the term has been or is actually used. But how?

(Flynn, 2011, p. 3)

What are human rights? At first glance this question may sound redundant if, as mentioned in the introduction, people seem to be familiar with this word, frequently used by the various media, both for foreign and domestic policy issues. Yet, the widespread use of a term does not always ensure a clear understanding of its meaning, making it entirely reasonable to ask: 'What do we mean by human rights?'. To address this, one solution may be to dissect the question into its two components, 'rights' and 'human', and determine whether the meaning is given simply by their sum or whether a further meaning emerges from their conceptual relationship. Another avenue is to define human rights by scrutinizing their multifaceted role in contemporary global affairs[65], while a third involves delving into the historical context of their inception. These approaches aren't mutually exclusive but rather complementary. Throughout this discussion, we will reference both historical and legislative milestones, highlighting their philosophical underpinnings. While part II will delve deeply into the historical roots, this part will primarily focus on the conceptual analysis of the terms. Thus, the starting question is: 'What do we mean by right and what do we mean by human?'.

While rights are frequently invoked in contemporary discussions of moral and political issues, the way people (and scholars) use this concept is not always clear and evident. The English word 'right' has two meanings: one refers to what is correct or just, and the other to people's

[65] "If we look at their role in 1948, when they were first officially introduced by the Universal Declaration of Human Rights (UDHR) and compare it with the role they play today in various contexts, one can easily realize that their impact has increased beyond anything the drafters in 1948 could have hoped or imagined. Human rights today are not merely the official normative language that most politicians speak around the world: they are standards that determine (a) the conditions of legitimate sovereignty; (b) whether the international community has a right to intervene (in different forms, up to military actions) against governments or power groups that violate them massively and systematically; (c) whether countries are eligible to enter the European Union; (d) the accountability of the forty-seven governments that are members of the Council of Europe (including Russia, Turkey, Hungary) to an international tribunal such as the European Court of Human Rights, which delivers binding judgements often leading to an alteration of national legislation; (e) the backbone and raison d'être of a number of international institutions (e.g., the UN High Commission) and influential NGOs; (f) a source of inspiration for millions of activists around the world, providing them with a common language and a shared basis of political initiative" (Caranti, 2019a, p. 4).

entitlement.[66] In this analysis we will focus on the second meaning,[67] although the semantic link between the two will often emerge and will be made explicit by analyzing other normative concepts.[68] Normative concepts guide individuals and institutions on how they should behave or what they should provide by offering standards for assessing the justice and appropriateness of actions and societal arrangements. Defining how rights are normative to us is essential to ensure they are consistently understood and applied within these evaluative frameworks.

Is such an analysis necessary, some might ask, given that the concept of right seems self-explanatory? Likewise, the concept of "human" may seem even easier to clarify: it is the biological term that denotes all species within the genus *Homo*, a group of hominid primate mammals that includes several extinct species such as *Homo neanderthalensis*. Among these species, *Homo sapiens* is the only one still existing today, and it is distinguished by several unique characteristics that set it apart from other living species, such as a highly complex and articulated symbolic language, enabling sophisticated communication and the expression of abstract concepts.[69] Therefore, a minimalist definition of a human being is usually as follow: "a member of the species *Homo sapiens*, characterized by advanced cognitive and linguistic abilities and complex social structures" (National Academy of Sciences US, 2010).

Can we thus provide a definition of human rights by combining these two concise and straightforward definitions, thereby specifying the concept of rights based on the concept of human? Are human rights, therefore, entitlements we possess solely by virtue of belonging to the human species? This is, in fact, the definition provided by Alan Gewirth, with the further

[66] Donnelly refers to the two meaning as, respectively, rectitude and entitlements (Donnelly, 2002, p. 7). I will also use "entitlements" as synonymous of this second meaning of "rights", rather than as a distinct concept as Joel Feinberg (1966) does. Feinberg emphasizes that entitlements highlight the status-value of rights—what individuals are owed or guaranteed, potentially independent of any active assertion—in contrast to the active element of claiming rights. While acknowledging the significance of claim-rights as the logical form of the concept of rights and the crucial distinction between "active demand" and "passive status" (as will be elaborated later on in the first chapter), I will still refer to entitlements as a general synonym for rights, similar to how the term is used in everyday language.

[67] Even if, as we will see in paragraph 1.2, this second meaning can be distinguished in four distinct forms.

[68] In the end, rights as entitlements possess normative authority because they are seen as (morally or legally) *right*, in the broader sense of complying with the normative frameworks of the moral community or the constitutions.

[69] This advanced linguistic capacity facilitates the transmission of information across generations, not only through biological inheritance but also through cultural transmission, which fosters the development of intricate societies and civilizations (Boyd & Richerson, 1988; Tomasello, 2009).

specification of what type of rights these are: "Human rights are a species of moral rights: they are moral rights that all persons have equally simply by being human" (Gewirth, 1982, p. 1). However, it is precisely when one seeks to delve deeper by adding specifications that the first challenges begin to arise. There is, in fact, no consensus among scholars regarding the categorization of human rights as moral rights, and some scholars prefer alternative terms such as natural, basic, or universal rights (Donnelly, 2002; Pogge, 1995; Shue, 1980). Are these terms synonymous, or do they each carry distinct meanings? Moreover, there are those who argue that human rights do not belong to any of these categories but reflect merely power relations in the international realm (Kennedy, 2004; Koskenniemi, 2001; Mutua, 2000).

Furthermore, even at the foundational level of the definition—human beings—there are questions that need answers: what makes human beings deserving of a specific category of rights? Is it their possession of certain capacities, characteristics, or fundamental goods (Nussbaum, 1997; Sen, 1984; Nickel, 2007; Finnis, 2011) or is it grounded in a single overarching-concept like human dignity (Nagel, 1986; Waldron, 2015; Buchanan, 2017; Caranti, 2019b)? If the latter, does dignity derive its normative force self-evidently, as articulated in the Universal Declaration of Human Rights and the various international conventions, or is it just a construction of positive law? In order to develop a constructivist response to all these questions, I am going to first clarify what we mean by the concept of right in general (chapter 1), distinguish it from other normative concepts (chapter 2), and finally explore how the condition of being human defines, or limits, the scope and application of human rights (chapter 3).

CHAPTER 1
THE CONCEPT OF RIGHT

Just as a right involves an authority to claim that to which one has a right, so also is moral obligation conceptually tied to what the moral community can demand (and what no one has a right not to do). And what I am morally free to do is what no one can justifiably demand that I not do either as a member of the moral community (no conflicting moral obligation) or as an individual (no conflicting right).

(Darwall, 2006, p. 20).

1.1. THE NATURE: OVERLAPPING BOUNDARIES

1.1.1 The Separation Thesis of Moral and Legal Rights

Rights permeate our society.[70] Most countries enshrine a list of entitlements within their constitutions, asserting, for instance, that the government shall not infringe upon the freedom of religion of its citizens. These entitlements, which we can refer to as 'institutional' or 'legal rights', constitute legal norms characterized by specific forms and functions (Herstein, 2023). However, the concept of rights extends significantly outside the legal sphere as well. When teenagers yell at

[70] While the concept of individual rights, as understood in Western liberal traditions, may not hold the same prominence in all societies, it seems nonetheless unproblematic to acknowledge that rights occupy a crucial position in most countries worldwide. One of the main theses of this work (which will be tackled fully in the final chapter) consists in the more contested claim that human rights are universal, and that they should be recognized by any reasonable person, regardless of their cultural framework. However, in this preliminary context, where we are discussing rights in general, it is far less controversial to assert that the concept of rights is widely acknowledged by the majority of the world's nations, even if the focus and the role they play in their broader moral framework is different.

their parents: "You have no right to enter my room without knocking!" they are not invoking a legal right but rather a different form of entitlement.[71] This request is not arbitrary; teenagers try to substantiate its legitimacy by appealing to a more fundamental right—the right to privacy—even if this specific right is not explicitly codified for this context in an actual law.[72] We can refer to this second type of rights as 'moral right',[73] an entailment not based on an existing legal framework but rather on a different normative system.[74]

Keeping in mind the distinction of morality and law emphasized in the introduction (section B), some key differences between the two types of right have already emerged. Laws function in a binary manner: an action is either legal or illegal (legal if it conforms to established rules and illegal if it violates them); conversely, morality, characterized by public accountability and not by contingent institutionalized regulations, does not adhere to such clear-cut distinctions, but encompass shades of gray: an action may be deemed moral or immoral, but it is also possible to rank different types of moral or immoral actions based on their proximity to or deviation from the normative principles of reference.[75] Given these numerous differences, it is legitimate to ask if

[71] I will revisit this example multiple times throughout the chapter, gradually adding layers of clarity and detail. Using examples involving children and teenagers in discussions about rights often sits at the intersection of various approaches and theories (N. MacCormick, 1982; O'Neill, 1992; Wenar, 2013), but I believe that this very complexity makes these examples particularly valuable for elucidating key concepts.

[72] And it is probably impractical for it to be so, similar to Amartya Sen's example of introducing a wife's right to be consulted in family decisions (Sen, 2004, p. 345). This is probably because, for the law to penetrate what has traditionally been the primary social unit—the family—stronger and more enforceable justifications must be provided. Nevertheless, just as a husband who never consults his wife on family decisions is committing a moral wrong, similarly, a father who constantly enters his daughter's room without knocking is wronging her. Even though privacy rights have been extensively legislated, it is unlikely that the law will go so far as to regulate such intimate family matters. Yet, the principle behind the right is (morally) rightly invoked by the teenager when she asks her father to stop entering the room without knocking.

[73] The other common term used in opposition to a legal right is that of natural right, which is also employed in this paragraph to differentiate types of rights, as distinct from forms and functions of rights (Wenar, 2005). We will return to this natural-moral differentiation in the whole part II. I will just briefly say here, without further argumentations, that from a constructivist perspective, saying that human rights are moral rights, rather than natural rights, means affirming not only that they are universal rights, but also that this universality is not given in a foundationalist manner, but in a proceduralist one.

[74] For now, I will stick with a minimal distinction like this, which is not only embraced from a constructivist perspective but can also be acknowledged by scholars who view morality as synonymous with customs. Even if devoid of its universal stance, there remains a significant difference between the normativity of a legal right within a codified system of law and the normativity of other types of customs and habits (like the norms of netiquette). The theme of universality will be gradually introduced, chapter by chapter, with a specific focus on part II.

[75] An action may even exceed the specific normative principle it refers to, thereby falling within the realm of what is considered supererogatory, either to a greater or lesser extent (Heyd, 2024).

there is any connection between the two concepts. This question lies at the heart of the long-standing debate between juspositivists and jusnaturalists.

According to juspositivism (or legal positivism),[76] law should not be interested in the content of morality, independently of their universality or not. Law should be a specific and independent normative system; this is the so-called *Separation Thesis.*[77] Instead, according to jusnaturalism (or natural law theory)[78], there is no clear division between the notion of law and the notion of morality. Though there are different versions of jusnaturalism, all subscribe to the thesis that there are at least some laws that depend for their "authority" not on some pre-existing human convention, but on the logical relationship in which they stand to moral standards. This is called the *Overlap Thesis*: the idea that the concepts of law and morality intersect in some way.[79]

---

[76] To avoid asymmetry in definitions, I will predominantly use the terms "juspositivism" and "jusnaturalism" to refer to legal positivism and natural law theory, respectively.

[77] Sometimes the separation thesis is understood in terms of two specific versions: the *Fallibility Thesis*, which emphasizes that legal systems can be morally fallible and that the correctness of legal rules does not necessarily align with moral truths, and the *Neutrality Thesis*, which advocates for the legal system's neutrality on moral issues, suggesting that the law should not endorse particular moral values and should allow individuals to follow their own moral convictions. In this work, I will refer to a conceptualization of the Separation Thesis that includes the Fallibility Thesis but excludes the Neutrality Thesis, as outlined by Forst. See the footnote 33 for further elaboration.

[78] Scholars sometimes distinguish two kinds of theory that go by the name of natural law, one regarding the moral field and another one regarding the political field. The first is a theory of morality according to which moral standards that govern human behavior are, in some sense, objectively derived from the nature of human beings and the nature of the world. The second is the one that has properly to do with laws and subscribe to the overlap thesis. As an empirical matter, many natural law moral theorists are also natural law legal theorists, but the two theories, strictly speaking, are logically independent. One can deny natural law theory of law but hold a natural law theory of morality. John Austin, the most influential of the early legal positivists, for example, denied the Overlap Thesis but held something that resembles a natural law ethical theory.

[79] For simplicity, I will for now concentrate only on the separation and overlapping theses as the contrasting and distinguishing features of jusnaturalism and juspositivism. However, it's important to acknowledge the vast array of diverse and sometimes conflicting theses associated inside each of these longstanding movements. Carlos Nino, for instance, attempted to delineate all the primary positivist theses, while simultaneously asserting that: "no positivist thinker defends all the positions listed above. Hardly any two authors who claim to be positivists support the same sub-set of theses among those which have been mentioned" (Nino, 1980, p. 520) Nino's list of positivist theses is as follow: (i) There are no prescriptive 'natural laws', that is, absolutely valid standards establishing rights and duties, which are applicable in every society and time and which can be inferred from the 'nature' of man or of human reason or from the general arrangement of the universe. (2) There is no rational way of objectively justifying the validity of moral standards and ideals about the rights and duties that men have or about the justice of social institutions. (3) The mere fact that a legal system, whatever the content of its rules, is in force in a certain society is a reason for considering it morally justified and for asserting that the community and officials have the moral duty to obey and to apply it. (4) Legal standards are distinguishable from other social standards and, in particular, from the moral standards the society accepts. (5) Legal systems are self-sufficient in providing solutions for any possible case; the law has no gaps, contradictions, linguistic vagueness or ambiguities, etc. (6) The opposite of (5), namely that legal system may not contain a determinate solution for some cases. (7) The law consists only of standards explicitly enacted by centralized

Thus, through the separation thesis, juspositivism endorses a primarily descriptive approach to law, whereas jusnaturalism, together with the overlap thesis, advances a normative approach in which law is grounded in moral truths. But which stance should a constructivist adopt? While a definitive answer to this question will only emerge at the conclusion of Part II, at this juncture we can begin to explore what a constructivist might argue concerning the separation and overlap theses to understand their view on the relationship between moral and legal rights. To facilitate this inquiry, let us expeditiously scrutinize the enduring and renowned debate between Hart and Dworkin concerning the nature of law: does it derive solely from social facts, or does it encompass moral precepts?

1.1.2 Juspositivism and The Rule of Recognition

The separation thesis was considered by H.L.A. Hart's precisely the core of all the juspotivists proposals: the validity of law does not hinge on its conformity to moral principles but rather on its recognition and acceptance within a particular legal system. One of Hart's first seminal works, aptly titled "Positivism and the Separation of Law and Morals" (1958), rigorously examines the foundational principles of legal positivism, notably challenging John Austin's Command Theory, which laid the foundation for analytical jurisprudence. According to Austin's theory (1832), laws were command of the sovereign backed up by sanction; this, however, insufficiently captured the psychological and normative elements of obligations that do not merely arise from the fear of punishment but also from the internal acceptance and recognition of legal rules within a community.[80] To surpass this perspective, Hart introduces the concept of rules as standards of

organs (i.e., statutes). (8) The law includes only those standards which are recognized by judges by reason of their originating in some events occurring in time and space. Standards which judges apply in their decisions only because their content seems just to them or because they are supposed to be of divine origin, are not part of the law. (9) A legal system in force in a certain society can be identified only by taking into account empirical facts, just as the judicial recognition of its standards, disregarding any consideration about its moral value or justice.

[80] Austin definition of law incorporate three crucial components: command, sanction, and sovereign. Hart challenges each aspect of this definition in his paper: "The situation which the simple trilogy of command, sanction, and sovereign avails to describe, if you take these notions at all precisely, is like that of a gunman saying to his victim, "Give me your money or your life." The only difference is that in the case of a legal system the gunman says it to a large number of people who are accustomed to the racket and habitually surrender to it. Law surely is not the gunman situation writ large, and legal order is surely not to be thus simply identified with compulsion" (Hart, 1958, p. 603). Hart starts his critique by noting that reducing the concept of law to a merely commands being enforced by threats falls short in fully explaining modern legal systems, that encompass laws governing various aspects such as contract formation, wills, marriages, and other executory instruments. Hart identifies these laws as "power-conferring rules,"

behavior, which entail both external (predictive) and internal (evaluative) aspects: while external statements describe legal efficacy, internal statements involve moral evaluations of legal norms.

Furthermore, for Hart, a system of law combines two types of social rules: primary rules and secondary rules (Hart, 1961, pp. 92–98). Primary rules impose obligations or duties, demanding conformity with significant social pressure on deviators[81]. Secondary rules grant powers to introduce, change, or modify primary rules; it is thus through the application of secondary rules that we find the primary rules. Hart delineates three specific types of secondary rules: the rule of change, the rule of adjudication and the rule of recognition, but is specifically this last one that serves as the ultimate criterion for legal validity, determining whether a rule conforms to legal requirements[82]. Accordingly, a rule is legally valid and therefore law, if it conforms to the requirements set forth in the rule of recognition, which in itself can neither be valid nor invalid: the proof of its existence is in the social fact of its acceptance.[83] The rule of recognition serves as

---

asserting that they are not merely commands backed by threats but rather rules that establish a framework within which individuals can define the extent of their rights, obligations, and liabilities. Moreover, Hart thinks that the word "command" implies a top-down stable hierarchy of men, with rules being purely other-regarding, something which is not true in modern legal systems, as legislations often have a self-binding force. Ultimately, Hart contests Austin's definition of "sovereign," rejecting the idea of habitual obedience and the notion of an untrammeled will. Mere habits of obedience cannot justify a new legislator's authority, or the continued validity of laws made by past legislators. Rules, not habits, are necessary for the continuity of laws, as rules entail not only behavioral convergence but also a shared attitude towards standards of behavior. Austin's suggestion that in democracies the electorate forms the sovereign, will leads to the paradoxical notion of society obeying itself, which contradicts the idea of sovereignty Austin was trying to implement.

[81] Society's primary rules are not necessarily its codified statutes. The foremost example is the common law.

[82] If a person is murdered, society turns to the court to ascertain which primary rules govern the situation, and the court determines whether the primary rules have been violated. In its decision-making process, the court may reference legislative enactments pertaining to murder. The court's authority to determine whether a primary rule has been breached establishes a secondary rule of adjudication, indicating that the court serves as the adjudicative body. This constitutes the application of rules of adjudication. Moreover, when the court selects the legislative enactment regarding murder as the governing primary rule, it involves the application of rules of recognition. Thus, two secondary rules—recognition and adjudication—are utilized in the process. Still, the most important secondary rule used in the murder example is clearly the rule of recognition: when the court refers to the laws against murder, it identifies the source of the primary rule. See however the next footnote on the difficulty of conceptualizing the rule of recognition as a secondary rule.

[83] Surprisingly though, Hart's elucidation on the fundamental structure of the rule of recognition was rather nebulous, as underlined by Scott J. Shapiro. Hart often depicted the rule of recognition as a criterion for determining the content of the law within a specific legal system. For instance, he described the British rule of recognition as "whatever the Queen in Parliament enacts is law." However, reducing the rule of recognition to merely a test of legality poses challenges within the Hartian framework which seems to recognize only two categories of legal rules: those that impose duties and those that confer powers. Nonetheless, the recognition test does not neatly align with either of these categories (Shapiro, 2009, p. 240).

the bedrock principle upon which the inquiry into the basis of legal rules concludes. In order to maintain a functional legal system, citizens are required to approach primary rules from an external perspective while officials responsible for identifying and applying primary rules through secondary rules must adopt an internal viewpoint. This internal perspective is indispensable as it serves as the sole means through which reasons or justifications for enforcing, creating, or altering obligations can be articulated. [84]

This theoretical apparatus provides a more robust foundation for understanding the separation thesis, allowing for the acknowledgment of "bad laws" that remain legally valid if they comply with the internal rules of the legal system. To explore these points, Hart discusses the viewpoints expressed by Gustav Radbruch and Hans Kelsen surrounding the Nazi regime's extreme corruption in the legal system (Hart, 1958, p. 616) and especially, in what became known as the Hart-Fuller debate, the decision by a German postwar court (the "Case of the Grudge Informer"). The case in question is the following: a German woman denounced her husband to the authorities in accordance with the anti-sedition laws of 1934 & 1938. He had made derogatory remarks about Hitler. The husband was prosecuted and convicted of slandering the Fuhrer, which carried the death penalty. Although sentenced to death he was not executed but was sent as a soldier to the Eastern front. He survived the war and upon his return instituted legal proceedings against his wife. The wife argued that she had not committed a crime because a court had sentenced her husband in accordance with the relevant law of the time. However, the wife was convicted of 'illegally depriving another of his freedom', a crime under the Penal Code, 1871, which had

[84] Joseph Raz challenges Hart's distinction between the internal and external aspects of rules, asserting that it forces individuals into a binary choice: either assume the detached stance of an observer assessing the legal system's effectiveness or embrace the internal role of endorsing the law's moral authority. How Raz introduces a third category: that of lawyers or law teachers elucidating legal principles to others. This allows for internal statements to be made without necessarily advocating them as normative standards. For instance, one might vehemently oppose capital punishment yet, within the framework of their country's legal system, write a legal opinion stating, "Given the law on this point, he ought to be hanged." In this context, the use of "ought" does not signify an endorsement of the rule's moral content (Raz, 1970, 1979) Neil MacCormick tries to resolve this apparent contradiction in Hart's theory of rules by arguing that the internality of a statement is determined by understanding, not will. It's possible to comprehend a norm, make judgments based on it, and still be against or indifferent towards it. MacCormick suggests further categorizing the internal point of view into two types: mere understanding of what the rule demands and a volitional commitment to upholding the rule. He concludes by positing the reasonable assumption that no legal system can function without at least some individuals who genuinely care about maintaining patterns of conduct and have a volitional commitment to the rule (N. MacCormick, 1977, 2003).

remained in force throughout the Nazi period. The court described the Nazi laws as "contrary to the sound conscience and sense of justice of all decent human beings" (Fuller, 1978, pp. 245–255).

Lon Fuller contended that in this case the court's decision was the best one; Hitler's regime was so harmful to morality, that there was nothing in the system that could qualify to be called a law as they did not comply with his desiderata. He stated that the Nazi laws lacked the necessary internal morality required in the process of law making, which gives laws respect and makes them obligatory to be followed by citizens. Fuller outlines eight principles that he believes constitute the inner morality of law, which must be adhered to for a legal system to function justly[85]. These principles are inherently based on moral considerations, and it's morally wrong to expect people to follow laws they cannot possibly obey or to administer laws in a manner inconsistent with their original intent. He contends that law must serve a higher purpose of achieving social order through moral scrutiny, suggesting that failure to comply with these moral principles renders laws invalid.

Hart instead acknowledges that both alternatives in this specific case are undesirable: let the woman go unpunished is not morally correct but punish the woman would set a dangerous precedent of retrospective law. Hart was strongly against the Court's decision to introduce a concept of morality and deciding the statute which protected the woman was no law at all but argued that the Nazi regime could have been considered an exceptional circumstance for the application of retrospective of laws, if the Courts were afraid that Hitler's accomplices would be acquitted. This is more broadly connected with questions of justiciability: while rules within a legal system are typically identifiable, there are occasions where they become indeterminate, requiring judicial discretion for resolution, leading to what is known as the 'problem of the core

[85] These formal principles encompass generality, specificity, prospectivity, intelligibility, consistency, feasibility, stability, and congruence (Fuller, 1978, pp. 229–246). They are considered formal as they pertain to the structure or form of the norms governing our behavior and hey work together as an integrated whole. A law that is clear but retroactive (violating prospectivity) or a that is consistent but impossible to comply with (violating feasibility) cannot guide behavior effectively. Fuller argued that it is morally wrong to expect citizens to obey laws they cannot possibly follow or to enforce laws inconsistently with their intended purpose. Law, in his view, must serve the higher purpose of promoting social order through moral scrutiny. In the case of the Grudge Informer, while the court's decision might also appear to violate the prospectivity criteria, the Nazi legal system had collapsed on nearly every dimension: the laws were secretive, contradictory, retroactive, and impossible to obey, making the moral and legal expectation of compliance unreasonable. By highlighting the systemic failure of these interlocking principles, Fuller shows that the court's retroactive judgment, while technically a violation of prospectivity, was justified given the complete absence of morality in the preceding regime.

and the penumbra' where the explicit content of a law may be clear (the core), but there are areas open to interpretation (the penumbra)[86].

Hart exemplifies this with the example of a legal rule that forbids you to take a vehicle into the public park. “Plainly this forbids an automobile, but what about bicycles, roller skates, toy automobiles? What about airplanes? Are these, as we say, to be called "vehicles" for the purpose of the rule or not? If we are to communicate with each other at all, and if, as in the most elementary form of law, we are to express our intentions that a certain type of behavior be regulated by rules, then the general words we use - like "vehicle" in the case I consider - must have some standard instance in which no doubts are felt about its application. There must be a core of settled meaning, but there will be, as well, a penumbra of debatable cases in which words are neither obviously applicable not obviously ruled out.” (Hart, 1958, p. 607). Judicial discretion arises from the inherent uncertainty in social guidance and the vagueness of words, what Hart called the “open texture”[87] of language: despite attempts to establish standards of conduct, interpretation is always required.[88] Hart views his theory of adjudication as a balancing act between two opposing viewpoints: strong formalism, which entails judges rigidly applying written rules without discretion, and extreme realism, where judges have unfettered freedom to decide cases at their discretion, especially in challenging circumstances.

---

[86] “We may call the problems which arise outside the hard core of standard instances or settled meaning "problems of the penumbra"; they are always with us whether in relation to such trivial things as the regulation of the use of the public park or in relation to the multidimensional generalities of a constitution. If a penumbra of uncertainty must surround all legal rules, then their application to specific cases in the penumbral area cannot be a matter of logical deduction, and so deductive reasoning, which for generations has been cherished as the very perfection of human reasoning, cannot serve as a model for what judges, or indeed anyone, should do in bringing particular cases under general rules. In this area men cannot live by deduction alone. And it follows that if legal arguments and legal decisions of penumbral questions are to be rational, their rationality must lie in something other than a logical relation to premises” (Hart, 1958, pp. 607–608).

[87] “In all fields of experience, not only that of rules, there is a limit, inherent in the nature of language, to the guidance which general language can provide. There will be plain cases constantly recurring in similar contexts to which general expressions were clearly applicable […] but there will also be cases where it is not clear whether they apply or not.”(Hart, 1961, p. 123).

[88] This consideration extends beyond law to encompass any system of rules, including games (Hart, 1961, pp. 109–114). Even within a game as structured as chess, which operates within a closed set of rules, there are instances that necessitate referee adjudication—such as determining what constitutes unsportsmanlike conduct. However, for Hart, the distinction between games and law lies precisely in their referees or judges: in law even a supreme court operates within a system of rules that are sufficiently determinative at the core, enabling the application of accurate standards to judicial decision-making (Hart, 1961, pp. 138–143).

Thus, for Hart the separation thesis does not contest the truistic claim that the development of law, at all times and places, has been profoundly influenced by conventional (or unconventional) morals. Still, this does not necessarily imply that a legal system must exhibit some specific conformity with morality or justice. The only final end the law should stick to is mere survival, without necessarily aiming for any broader moral objectives: this is what Hart refers to as the minimum content of the natural law.[89] While other contingent moral principles may fluctuate, positive law can remain valid without any inherent dependence on the broader framework of morality. This is precisely the point that Ronald Dworkin aims to contest, by attacking the use that Hart makes of the rule of recognition and, in the end, by rejecting the separation thesis.

1.1.3 Dworkin Constructivist Conception of Law as Integrity

Dworkin starts his first attack in "The Model of Rules I" by setting out three theses to which he believes Hart and most legal positivists are committed: a) Law is a set of rules identified by specific criteria related to their pedigree or adoption process, rather than their content. These criteria help distinguish valid legal rules from spurious ones and from other social rules, such as moral rules; b) The set of valid legal rules is exhaustive, meaning if a case isn't clearly covered by such a rule, it can't be decided by simply "applying the law." Instead, officials like judges must exercise discretion, either by creating new legal rules or supplementing existing ones; c) Having a "legal obligation" means falling under a valid legal rule requiring specific actions or forbearances. Legal rights, powers, privileges, and immunities imply corresponding obligations on others. When a judge exercises discretion, they aren't enforcing a legal obligation on that issue because no valid legal rule applies.[90]

---

[89] Hart is criticizing a specific strand of the natural law tradition, which maintains that there exists a distinct moral system consisting of principles discernible through reason. According to this view, any law must conform to these principles to be considered legitimate. This perspective, as defined by Hart, is grounded in the "teleological view of nature," which posits that everything, including humans, inherently moves towards a specific end, or telos, and in a jusnaturalist perspective, laws are expected to facilitate the attainment of this telos through their commands. However, a constructivist position diverges from this stance by not asserting any substantive teleological claim, but simply describing a proceduralist approach to deliberating about final ends. This distinction will be revisited in part II, especially chapter 5.

[90] "(a) The law of a community is a set of special rules used by the community directly or indirectly for the purpose of determining which behavior will be punished or coerced by the public power. These special rules can be identified and distinguished by specific criteria, by tests having to do not with their content but with their pedigree or

The first principle, also called pedigree thesis, which affirms that exists a master rule for distinguishing law from non-law, is the main core of the dispute between Dworkin and Hart. What Dworkin is doing here is reinterpreting Hart's rule of recognition as follows: a master rules that specify essential features.[91] Possession of some features by a proposed primary rule is regarded, by the secondary rule of recognition, as a conclusive evidence that the rule is a valid primary rule of law; instead, the duty to follow the secondary rule, and consequently the primary rule, hinges upon their general acceptance. A classic illustration of this acceptance, often cited by positivists, is the majority of officials consulting enactments of Congress to resolve legal matters. This practice establishes a duty for all officials to follow suit.

Dworkin argues that this explanation fails to address why courts have a duty to follow the legislature. For instance, how does one justify a judge's decision to distribute an estate according to the statute of descent as in Riggs v. Palmer?[92] Hart would assert that the secondary rules of recognition identify the sources of law, such as legislative enactments or judicial precedent, which then pinpoint the primary rule that ultimately governs in the specific case at hand. According to Dworkin, this reconstruction misses the true normative core of law, which relies on morality: legal

---

the manner in which they were adopted or developed. These tests of pedigree can be used to distinguish valid legal rules from spurious legal rules (rules which lawyers and litigants wrongly argue are riles of law) and also from other sorts of social rules (generally lumped together as "moral rules") that the community follows but does not enforce through public power. (b) The set of these valid legal rules is exhaustive of "the law," so that if someone's case is not clearly covered by such a rule (because there is none that seems appropriate, or those that seem appropriate are vague, or for some other reason) then that case cannot be decided by "applying the law." It must be decided by some official, like a judge, "exercising his discretion," which means reaching beyond the law for some other sort of standard to guide him in manufacturing a fresh legal rule or supplementing an old one. (c) To say that someone has a "legal obligation" is to say that his case falls under a valid legal rule that requires him to do or to forbear from doing something. (To say he has a legal right, or has a legal power of some sort, or a legal privilege or immunity, is to assert, in a shorthand way, that others have actual or hypothetical legal obligations to act or not to act in certain ways touching him.) In the absence of such a valid legal rule there is no legal obligation; it follows that when the judge decides an issue by exercising his discretion, he is not enforcing a legal obligation as to that issue." (Dworkin, 1967, pp. 17–18)

[91] One may raise doubts as to whether this assertion holds true. For instance, Hart nowhere imposes a pedigree requirement on the rule of recognition; indeed, in certain instances, he explicitly allows that the criteria of legality may reference moral considerations. See Hart (1961, p. 204) and Shapiro (2007, p. 8).

[92] In this case, a man named Francis Palmer changed his will to leave his estate to his grandson, Elmer Palmer. However, before Francis Palmer passed away, his grandson Elmer murdered him in order to inherit the estate sooner. After Francis Palmer's death, Elmer was sent to prison for murder, yet he still inherited the estate according to the will. This outcome was based on the legal principle that a person should not be punished twice for the same crime. Palmer's other relatives contested the will, arguing that Elmer should not inherit because he had killed Francis to receive the inheritance. The case went to court, and ultimately, the court ruled in favor of the other relatives. The court decided that it would be against public policy and morality to allow Elmer to profit from his crime by inheriting the estate.

principles can be binding on judges simply because of their intrinsic moral properties and not because of their pedigree.[93]

In Dworkin's most refined critique of positivism, encapsulated in "The Empire of Law," the central focus lies on legal interpretation and the possibility of theoretical disagreements among judges.[94] He argues that legal practice fundamentally revolves around argumentation (which entails participants advancing diverse claims regarding the demands of the law and substantiating these claims by providing reasons) and contends that legal interpretation is, at bottom, "constructive". Constructive interpretation is the process of "imposing purpose on an object or practice in order to make it the best possible example of the form or genre to which it is taken to belong." (Dworkin, 1997, p. 52). A purpose makes an object the best that it can be when it both "fits" and "justifies" the object better than any rival purpose. A purpose "fits" the object to the extent that it recommends that the object exists or that it has the properties it has. A purpose is "justified" to the extent that it is a purpose worth pursuing. The constructivist interpretation must first impute a point to the particular practice that presents it in its best light, namely, one that best fits and morally justifies it. Then, she must use this point to ascertain the grounds of law for the particular system.

By reframing the dispute in these terms, Hart and the positivist argue that judges should apply the law as it is, focusing on the rules and precedents established by the legal system, avoiding making moral judgments; in this sense judges must sometimes exercise strong discretion because he takes the law to consist in those standards socially designated as authoritative. Conversely, Dworkin believes that judges do not have strong discretion precisely because he denies the centrality of social guidance to determining the existence or content of legal rules. Dworkin tries to show that the law contains norms that are binding even though they have not been the subject of past social guidance: they are binding, rather, because of their moral content. He proposes the idea of "law as integrity," suggesting that legal principles should be derived from a coherent moral framework that respects individual rights and justice; even in hard cases where legal principles

[93] To see how Hart and his followers answer to these critique and thus have a clearer picture on the whole Hart-Dworkin debate, see Shapiro (2007).

[94] This represents a shift away from the emphasis on the distinction between rules and principles outlined in *The Model of Rules I*, which was no longer regarded as the contentious crux of positivism.

seem to conflict or provide no clear answer, judges must look beyond the rules and precedents to interpreting the underlying moral principles of the legal system. Even when social guidance runs out, the law does not, for moral guidance does not.

This approach accommodates theoretical interpretations in law, as disagreements stem from differing views on law's moral value and its relation to practice. Consequently, genuine moral disagreements lead to genuine legal disagreements, unlike the standard plain-fact view adopted by positivist,[95] which may dismiss theoretical disagreements as insincere or incoherent. Instead, the disagreements are engaged continuously by practically asking: *"*which interpretation best fits and justifies the legal practice?". Judges that are disciplined by the principle of integrity must render decisions consistent with precedent and the structural design of the legal system, base their rulings on principle rather than policy or political expediency, and apply relevant principles consistently across similar cases. For example, a Marxist judge could not simply impose a radical redistributive conception of equality; under law as integrity, their reasoning must respect constitutional coherence and existing legal principles, even while advancing morally justified outcomes. In this way, Dworkin shows that legal reasoning is an open-ended, morally guided interpretive process: disagreements are ongoing and inherent to the law, and their persistence is not a flaw but a feature of a system that integrates legal and moral reasoning. Legal reasoning is not merely mechanical application of socially recognized rules but a principled morally guided interpretive process.

To summarize, the Hart-Dworkin debate centers around the rule of recognition, specifically whether judges should exclusively adhere to established legal rules and precedents (as Hart argues) or also consider broader moral principles in their decision-making (as Dworkin contends), especially in hard cases where the law is ambiguous or incomplete. Dworkin's constructivist proposal of "law as integrity" consists in showing how judges are morally disciplined: decisions

[95] According to Dworkin, the "plain fact" perspective comprises two fundamental principles. Firstly, it asserts that the foundations of law within any community are established by consensus among legal authorities; in essence, if these authorities concur that certain facts, denoted as type "f," constitute legal grounds within their system, then indeed, these facts are deemed as such. Secondly, it contends that only factual circumstances grounded in plain historical events may serve as legal foundations. "The law is construed solely as a reflection of past decisions made by legal institutions such as legislatures, city councils, and courts. For instance, if such a body has ruled that workers can seek compensation for injuries caused by fellow workers, then this ruling constitutes the law. Conversely, if the decision has been made differently, then that becomes the law. Consequently, questions of law can invariably be resolved by referring to the records where institutional decisions are documented" (Dworkin, 1997, p. 6).

must be consistent with precedent, ensuring that the law evolves coherently while remaining anchored in moral reasoning.

While some scholars have argued that this debate no longer warrants the scholarly and pedagogical prominence it has enjoyed over the past four decades (claiming that Hart clearly prevailed)[96] the core issue of the debate remains significant. As Shapiro emphasizes, the essential question was not about determining a victor but about understanding if law is "ultimately grounded in social facts alone, or do moral facts also determine the existence and content of the law?" (Shapiro, 2007, p. 54). Accordingly, any proposal that aspires to be genuinely constructivist needs to side with Dworkin in this debate and reject the separation thesis in favor of some sort of overlap between law and morality

This core question is less abstract and more urgent than it might initially seem: we continuously question whether a particular law is just or not and this is not an open question in the Moorean sense but a practical necessity that requires an answer.[97] The constructivist answer is that we can conceive of unjust laws in a non-contradicting way because some social contingencies can lead to the institution of some sort of laws that go against one or more moral principles, and in this case is the judges that needs to review the law in light of moral considerations to ensure coherence and justice in legal decisions. However, while it's reasonable to criticize unjust laws by distinguishing between moral and legal principles, the same logic does not transfer as straightforwardly to rights: laws seem to impose obligations and restrictions on behavior, while rights are mostly perceived as protective measures that uphold justice and prevent abuse. Consequently, the question arises as to how a right can be considered unjust or immoral. This discrepancy may stem from the overarching logical form of rights: being a claim.

[96] "The moment now seems opportune to step back and ask whether the Hart/Dworkin debate deserves to play the same organizing role in the jurisprudential curriculum of the twenty-first century that it played at the close of the twentieth. I am inclined to answer that question in the negative, though not, to be sure, because I can envision a jurisprudential future without Hart's masterful work at its center. Rather, it seems to me – and, I venture, many others by now – that on the particulars of the Hart/Dworkin debate, there has been a clear victor, so much so that even the heuristic value of the Dworkinian criticisms of Hart may now be in doubt."(Leiter, 2003, p. 18).

[97] When I say that the question 'Is this law just?' is not an open question in Moore's philosophical sense, I mean that that the concept of justice can be explained and clarified not through intuition, as Moore suggests, but through practical deliberation. We will return to Moore's open question argument in more detail in Chapter 2, Section 2.3.

1.2 THE FORM: CLAIMABILITY AND CORRELATIVITY

1.2.1 Claim and Liberty rights

What, then, is the unifying link between legal and moral rights? Why do we employ the term "right" to describe both, despite their distinctiveness? Is there a logical form embedded within the notion of "right"? Indeed, there are several. Wesley Newcomb Hohfeld's seminal analysis, which remains one of the most pivotal and influential frameworks for categorizing rights, [98] delineates four jural equivalences or correlatives that underscore the significance of rights: claim, privilege, power, and immunity (Hohfeld, 1920). A claim-right entails a valid claim to a performance or forbearance, while a privilege signifies the absence of an obligation or duty; a power denotes a specific authority to modify the existing scheme of rights and duties creating liabilities, whereas immunity represents protection from such authority and the absence of liabilities[99]. For instance, consider the assertion "Daniel has a right to give his house to Lukas".[100] This statement can convey four distinct propositions, each imbued with its own meaning:

[98] For those delving into the philosophy of rights, Hohfeld (1879–1918) occupies a position similar to that of Gottlob Frege (1848–1925) in the philosophy of language. Just as Frege's ideas are foundational for anyone exploring the nature of language, Hohfeld's framework is regarded as an essential starting point for understanding the nature of rights. His analytical approach has become so influential that virtually every seminal work on theories of rights begins with a discussion of Hohfeld's work (Thomson, 1990; Waldron, 1984).

[99] It is generally agreed by rights theorists that in speaking of a particular "right," people often, if not always, have a combination of these forms in mind. Wellman (1985) argues that there is always both a core and one or more peripheral Hohfeldian relations constitutive of a given right. For instance, in speaking of a "right of free speech," someone may be assuming— without putting things in so many words—that he has a liberty (core) against all parties to speak his mind together with a claim against all others to their not preventing him from doing so. (Gilbert, 2018, p. 24). Similarly, the right of self-defense seems to embody both the liberty to use violence in defense and the claim-right against interference with this liberty. The notion of being "at liberty" to do something suggests not only the absence of obligation but also the expectation that others refrain from interfering. In discussions of rights, we can characterize these aggregation of various rights as "cluster rights," a concept advanced by Judith Jarvis Thomson (1990), who identifies them as a distinct category of rights, augmenting Hohfeld's list. Leif Wenar (2005) instead, emphasizes the foundational concept of freedom that resonates across Hohfeld's four distinctions. See the next footnote.

[100] Property rights are a prime example of 'molecular rights,' which are rights composed of more than one Hohfeldian element, encompassing all four essential components. My ownership of a house includes a power-right that allows me to waive the associated claim-right preventing others from entering the house. I can also transfer the entire set of molecular rights by selling the property, meaning these second-order rights can change or be extinguished through this transfer. Additionally, my property right grants me a privilege-right to enter and reside in the house.

1. Daniel has a legitimate request that others do not interfere with him giving his house to Lukas: he has a claim-right against others that they do not obstruct him in giving his house to Lukas.
2. Daniel has no obligation or duty not to give his house to Lucas: he has the liberty-right (privilege) to give his house to Lukas.
3. Daniel has the authority to give his house to Lukas: he has the power-right to change the existing scheme of rights and duties by giving his house to Lukas, thereby transferring her property rights in the house to him.
4. Some other entity (e.g., the state or Daniel's parents) lacks the authority to prevent Daniel from giving his house to Lukas: he enjoys a certain immunity-right regarding others' power to alter the existing scheme of rights and duties.

However, claim-rights are often considered, even by Hohfeld, as the quintessential meaning of "right" in its strict sense (Hohfeld, 1920, p. 36).[101] Thus, in most contexts, rights are most typically associated either with the form of claim-rights, or with (privilege) liberty rights, which is supposed to be the correlative and rivalry view.[102] A claim right is the entitlement held by an individual against another person or entity that carries a corresponding obligation owed to the right holder. For instance, returning to the example of teenagers requesting privacy, the right they assert is a

---

Lastly, it includes an immunity-right that protects my ownership by preventing others from altering any of the associated rights, including my privilege-right, claim-right, and power-right. Other classical molecular rights examples are religious freedom and patent rights. For a theory that interprets human rights as molecular rights, see Wellman (2011b).

[101] Joel Feinberg (1980, p. 155) is usually credited with generalizing this Hohfeldian assumption, significantly broadening its influence. For example, both Henry Shue (1980, pp. 14–15) and Jack Donnelly (2002, p. 8), two prominent human rights theorists, reference Feinberg when emphasizing the critical role of claiming in the framework of human right.

[102] The other two terms in Hohfeld's analysis, powers and immunities, are mostly considered as second-order liberties and claims, respectively. Powers are liberty rights that pertain to the modification of first-order rights. For example, the U.S. Congress has certain powers to modify the legal rights of U.S. citizens, as it can impose or remove legal duties. Conversely, immunities are claim rights that concern the modification of first-order rights. For instance, U.S. citizens have certain immunities, as outlined in their Constitution, which limit the positive powers of the U.S. Congress to alter their legal rights. As such, immunities and powers are often grouped together into "active rights" (liberties and powers) and "passive rights" (claims and immunities) or simply subsumed within claims and liberties by later authors (Freedman, 2014). Wenar for example, by focusing on the foundational concept of freedom, distinguishing the group of the active incidents—the privilege and the power— which entitle their holders to *freedom to* act in certain ways, and the passive incidents—the claim and the immunity—which entitle their holders to *freedom from* undesirable actions or states (Wenar, 2023).

claim right directed at their parents, who have a corresponding duty to respect the demand.[103] Liberty rights, on the other hand, exist without any corresponding duties to refrain from certain desired activities, essentially encompassing actions one is not prohibited from undertaking. I might assert a liberty right to spend my vacations lounging on a picturesque beach in Sardinia. Regrettably, no one is obligated to facilitate this specific exercise of my liberty right. I am simply affirming that I am free to go on vacation in Sardinia if I want to, and no one have a say on that. In logical terms: A person's liberty right to X entails their freedom to **Φ** or possess X, while a person's claim right to X implies an obligation on others to permit or facilitate their ability to **Φ** or have X.[104]

As an example, asserting a liberty right to free speech means asserting the freedom to speak without wrongdoing, but this liberty right does not automatically require others to assist in communication or prohibit them from preventing free speech. To assert such obligations would imply a claim-right to free speech. In contrast to claim rights, liberty rights are thus predominantly negative in nature.[105] To illustrate further: a world with only liberty rights, devoid of any claim

[103] This type of request also highlights various internal distinctions, such as those between negative and positive claim rights, as well as between *in personam* and *ad rem* claim rights. A negative claim right refers to the entitlement one holds to a particular good or service, which another party is obligated to provide. In contrast, positive claim rights are those held against others' interference in one's life or property: a teenager's claim to privacy represents a negative claim right against their parents, while their right to education constitutes a positive claim right toward the state, specifically directed at the Minister of Education. Furthermore, rights held *in personam* pertain to entitlements directed against a specifically identified duty holder, such as an educational authority, while rights held ad rem are not tied to any particular individual but rather apply universally, against every person (A. Ross, 2019, pp. 230–231). A teenager's right to education would lack practical significance if it were not enforceable against a specific, relevant, and competent authority, while the efficacy of their right to privacy would be significantly diminished if it did not extend to all potential violators of it (Jones, 1994).

[104] As Leif Wenar as shown, these implications hold often, not always, precisely because each of these forms of rights-assertion can also indicate a different "higher-order Hohfeldian incident": we have not only privileges and claims, but rights to alter our privileges and claims (power) and rights that our privileges and claims not be altered (immunity). Thus a more modest definition of liberty and claim rights is that: "A has a right to **Φ**" *often* implies a privilege, and "A has a right that B **Φ**" *often* implies a claim (Wenar, 2005, p. 230).

[105] Refer to the previous footnote for the distinction between negative and positive rights, and to the next one for a deeper understanding of the interconnections between negative rights and liberty rights. Furthermore, as we will explore in the following paragraphs, particularly through an examination of O'Neill's insights, positive rights present fewer challenges both theoretically and practically (O'Neill, 2016, pp. 35–42). That negative rights hold less theoretical weight than positive obligations is also underscored by Pogge, who has recently redefined our responsibilities towards the world's poor as a negative duty rather than establishing a positive rights with a correlative positive duty, arguing against perpetuating "unjust institutional arrangements" imposed on impoverished populations (Pogge, 2014, p. 20). However, when proposing specific measures to address these issues, questions arise regarding what upholding human rights entails for reforming the global economic order, as highlighted from Robert Howse and Ruti Teitel (2010).

rights, would essentially permit everything, with no acts or omissions being prohibited and no rightful claims of wrongdoing or neglect. Conversely, a world with solely claim rights and no liberty rights would mean everything is either obligatory or prohibited, leaving no room for mere permission.[106] This demonstrates once again that it is difficult to discuss liberty rights without also addressing claim rights, since the rigid interdependence of their logical forms makes it challenging to consider one without the other (Feinberg, 1970, p. 148).

Overall, the debate on claim rights and liberty rights reflects different perspectives on the role of rights in society, the relationship between individuals and the state, and the balance between individual autonomy and social responsibility; claim rights are often seen as essential for addressing social inequalities and promoting distributive justice, while liberty rights are viewed as safeguarding individual freedoms in a libertarian sense, preventing tyranny or authoritarianism.[107] This can lead to some misconceptions, such as the belief that utilitarians, who focus on maximizing the total sum of utility, rely primarily on claim rights, while Kantians, who emphasize the separateness of persons, rely primarily on liberty rights. It is important to clarify this misconception before moving on.

As previously noted, some philosophers that identify themselves as Kantian constructivists underline the deontological element in opposition to a teleological one.[108] Although this distinction

[106] This is a reimagining of a hypothetical world inspired by Joel Feinberg's example of Nowheresville, which is directly addressed in the following paragraph on the correlativity of duties and rights. See especially footnote n.143.

[107] It is important to emphasize, within the scope of this project, that human rights were initially developed primarily as negative rights against governments and other individuals, focusing on liberty rights such as civil and political rights. These rights aimed to ensure that people would be left alone, free from attack and interference. However, shortly afterward, the recognition of positive rights to basic necessities was addressed through the introduction of social and economic rights as human rights. For an analysis of all these generations of human rights, see Nickel and Etinson (2024).

[108] Some moral constructivists often reject this dichotomy, especially regarding the depiction of Kant as the central figure in deontology. The reinterpretations proposed especially by O'Neill (1989) and Herman (1993), have shown that Kantian ethics cannot be simply reduced to deontology. This is not only because it includes a significant component, identifiable in the Metaphysics of Morals, that can be aptly defined as virtue ethics, but also because the theory of practical reasoning generated by the categorical imperative is not deontological-deductive but rather "constructive". In this sense constructivist is precisely an alternative to the unfruitful debate between deontologist and utilitarians (Bagnoli, 2021b). However, regarding political constructivism especially, there is a notable tendency to retain the deontological inheritance. For instance, Forst (2002, p. 175) emphasizes that a constructivist theory of justice maintains its deontological character, as distinguished by Rawls (1971, 30) through the priority of the right, only if it rests on practical reason. Forst's analysis expands on Habermas's distinction of four essential characteristics of Kantian-procedural theory, the first being precisely its deontological nature character which "is aimed at the categorical normative validity of norms that can be universally justified with shared reasons" (Forst, 2002, p. 177).

was criticized by many scholars,[109] the actual debate that happened between scholars identifying as deontologists and those as consequentialists—particularly utilitarians—highlighted a key issue: how consequentialists often failed to adequately account for the separateness of individuals. In opposition, deontologists placed individual rights at the center of ethical and political discussions. Conversely, consequentialists typically view individual rights as secondary principles, justifiable only insofar as they contribute to creating greater overall good consequences. Deontologists argue that justice primarily concerns the relationships between individuals, rather than between individuals and states of affairs or abstract goods: each person is regarded as a distinct individual whose fundamental good should not be sacrificed for aggregate or universal good. They thus object that consequentialism maximization of the overall good fails to account for the types or quality of relations among individuals pursuing this good, as well as how the good is distributed among them.[110]

Other constructivists maintain this element without labeling it as deontology, referring to it instead as non-consequentialist (Richardson, 1990, p. 309). See footnote n.190 for further elaboration on this final point.

[109] As Timmermann (2015, p. 7) noted, it was C. D. Broad who introduced the term deontology, primarily defining it in opposition to teleological theories (of which consequentialism is merely a subset). However, this usage of "deontology" can obscure the distinctive and valuable aspects of different moral theories, leading to misguided questions while overlooking more important ones. For these reasons, some scholars suggest abandoning the term altogether or, at the very least, using it more precisely to refer to a specific group of 20th-century British ethicists, who sought to distance themselves from teleological theories. See also the previous footnote for some of the different positions of constructivists on this topic.

[110] Samuel Freeman (2011) have examined three types of consequentialist arguments for basic individual rights in response to these types of criticism, arguing that each has serious, if not insurmountable, problems. The first is the indirect consequentialist framework proposed by J.S. Mill, the second involves distribution-sensitive accounts of well-being and other goods (such as relied on by Philip Pettit, Larry Temkin, and Bill Talbott), and the third directly incorporates rights and other moral concepts into the good to be maximized. Firstly, Freeman argued that even if Mill has shown that basic rights and liberties are necessary for individual well-being, it doesn't follow that equal rights and freedoms are always necessary for maximizing individual well-being and thus he cannot properly deflect the critics that accuse consequentialist of failing to respect the separateness and equality of persons. Secondly, he applied similar criticisms to arguments advocating for the equal distribution of goods: incorporating into the consequentialist maximization equality and other distribution-sensitive values still does not make the case for equal basic rights and liberties. Lasty, he also challenged the third position represented by Amartya Sen, which argues that equal rights and fair distributions are intrinsic goods to be promoted for their own sake. He suggests that this position is not purely consequentialist but rather a pluralist intuitionist conception that involves balancing aggregate goodness against moral principles of fairness and individual rights. In the end, thus, the only suitable right-consequentialist approach needs to abandon the act-consequentialist principle, which assesses the rightness of actions solely by their individual tendency to maximize the good, and switch to an indirect-rule-consequentialist conception and defense of equal rights. "Taking maximally good consequences into account is rarely, if ever, a legitimate reason for individuals or even governments to act upon. Instead, it enters into rule consequentialist justifications (if at all), not at the level of deciding rightness of actions, but in the justification of the entire system of rules and specification of the kinds of rights that individuals ought to have. My argument does not address these positions, which on the whole seem to be a more promising way for consequentialists to argue for equal moral rights. (Freeman, 2011, p. 128).

The deontological critique therefore is not only that utilitarianism cannot convincingly employ the concept of claim rights, but that it cannot use the concept of rights in general without undermining it. Furthermore, deontologists, or constructivists in general, reject the accusation of relying solely on the liberty rights of the individual "unencumbered agent" (Sandel, 1982, p. 157).[111] They are instead primarily concerned with claim rights, as their focus is on the practical conditions of action (Bagnoli, 2022a, p. 4). And to understand what actions to take, it is crucial to comprehend what is prohibited, including the duties and obligations one has regarding others' rights.

1.2.2 The Correlativity of Rights and Duties

If the crucial notion of a right is that of a claim right, its core is further clarified through its correlation with the concept of duty, a requirement to perform or refrain from certain actions (Feinberg, 1970, p. 244).[112] This fundamental connection serves as the central foundation for understanding the meaning of rights, not only for constructivists but also for the majority of scholars, who seek to elucidate rights through their correlatives.[113] The correlation of duties *to* specific people and rights *against* specific people can be expressed as follow: if B has a duty to A, one can infer that A, as the beneficiary of that duty, possesses a claim right, if 'right' is understood

---

[111] In his critique of Rawls, Sandel not only attacks his conception of the person but also argues that every deontological theory of morality presupposes this implausible condition of the self, and thus every deontological conception of morality must fail: "Justice cannot be primary in the deontological sense, because, we cannot coherently regard ourselves as the kind of beings the deontological ethic—whether Kantian or Rawlsian—require us to be" (Sandel, 1982, p. 14). However, as shown by Forst, Sandel misses the point of his critique: "Kant explicitly distinguishes moral autonomy—acting according to general, self-given, universalizable laws—from the freedom of choice and action of legal subjects whose external relations are regulated by law. Sandel, however, misunderstands freedom from empirical determinations in the characterization of moral autonomy […] The issue is thus a completely different one. Kant is concerned with the morally right, Sandel with the ethically good […] Sandel thus connects an ethical thesis about choosing the personally good—a thesis he ascribes to Kant—with the assumption of a resulting normative consequence for the regulation of freedom of choice (i.e., the freedom to choose the good) —which is basically an assumption concerning relations in law; and this he does in order to criticize Kant's theory of morality. Moral autonomy is therefore misunderstood at times as ethical choice and at other times as the freedom of legal subjects" (Forst, 2002, p. 12).

[112] "A duty is what I must or must not do; in other terms, it is a matter of requirement" (Feinberg 1970, 244).

[113] This tradition dates back at least to Pufendorf, as Henry Richardson underline (2018, p. 77).

as the Hohfeldian correlative of duty, often referred to as a 'directed duty'.[114] Most contemporary rights theorists have precisely assumed that the best way to understand rights-claims is to clarify what it is a directed duty, and if it includes or not a more basic concept of 'plain duty', a duty independent of the correlation with a right.[115]

Judith Thomson, for example, tries to explain this difficulty by differentiating between doing something wrong (violating a general obligation) and wronging someone (violating a directed duty). Directed duties are distinctive because they establish a personal moral connection between the agent and the claim-holder: if the agent cannot fulfill the duty, they must either seek permission in advance or make restitution afterward (Thomson, 1990, pp. 200–202)[116]. For example, if you break a promise to your friend, you not only act wrongly but also specifically wrong your friend, since the duty was owed to them, and you are obligated to compensate for the harm caused. This relational element is central to Thomson's critique of moral theories, which often overlook the interpersonal nature of directed duties, like consequentialism which reduces all duties to general considerations of maximizing good outcomes (Thomson, 1990, pp. 124–148). Direct duties are thus understood by Thomson as situational applications or specific extensions of plain duties, but they are not universally obligatory, as not all plain duties give rise to direct duties in every context (Thomson, 1971).[117]

---

[114] This assumption regarding the existence of such rights remains unaffected by whether the right-holders harbor conscious interests or can assert claims on others (topics we will delve into in the paragraph 1.3.1); it relies solely on the presence of a duty towards the (putative) right-holder. On why duties are not rights, see Cruft (2006).

[115] "While operating with the assumption that a directed duty is constituted in part by a plain duty, rights theorists tend to assume that the duties or requirements in question are either institutional or moral. There is on this view no pertinent third category of duties, or, indeed, of claims" (Gilbert, 2018, p. 83).

[116] "What is it to have a claim? X's having a claim against Y is equivalent to Y's behavior's being constrained in a certain way. What way? […] There is, first, that, other things being equal, Y ought not just, all simply, infringe the claim even if Y would produce great good or advantage by infringing it: other things being equal, Y ought to seek a release from X in advance before doing what would, without a release, be an infringement of X's claim. There is, second, that, other things being equal, Y ought to compensate X later for such harms or losses as Y causes X by infringing the claim if Y infringes it. Seeking a release in advance, and compensating later, are acts like any others, and more will therefore be relevant to the question whether Y ought to do these things than merely that X has the claim. The presence of the claim shows itself in the fact that there has to be good reason for not doing these things if it is to be permissible to not do them." (Thomson, 1990, pp. 200–201).

[117] This distinction is illustrated in Thomson's famous violinist example, used in her defense of abortion, where the plain duty to preserve life does not necessarily impose a direct duty to sustain another's life under all circumstances (Thomson, 1971, pp. 48–49).

Contemporary rights theorists who, unlike Thomson, think that there is something special and distinct about duties toward a person, will still generally agree that such duties are not, or not simply, duties concerning that person; that is, they are not simply duties to do something to, for, or with that person.[118] These theorists tend to assume, rather, that the direction of a duty is determined by some external relational fact about it—irrespective of its content, and thus, that some duties arise from general principles that do not correspond to specific individual rights (Darwall, 2013c).[119] This, still leaves open other questions related to the directness of this duty and how it affects the way we think, act, and hold others accountable (Jonker, 2020). These inquiries will be partially revisited in Section 1.3, where the implications for the function of rights will be explored. Here I am just interested in underscore that the logical form of claim-rights implies also the logical correlation of (claim)-rights with (directed) duties, but not that all duties have correlative rights. A comprehensive understanding of the content and distribution of duties associated with a right is essential for its effective integration into the broader framework of personal and political behavior. As Henry Shue has contended, a claim of right remains incomplete until we have delineated, at least in part, what it would entail for a particular right to be realized and exercised (Shue, 1980).

Constructivists also endorse this assumption; yet, one of the most important and widely recognized exponent of the constructivist movement,[120] Onora O'Neill, has maintained even a stronger position: a moral theory that focuses on action should "see obligations or duties rather

[118] "Most rights theorists, as noted, do not see a duty toward X as, simply, a duty that concerns X, or a duty that concerns X in some particular way. Rather they accept the duty-plus assumption, amplified thus: directed duties are plain duties distinguished not by their content but by certain relations in which they stand. In particular, they involve a specific relation of the duty to the right- holder. As we shall see, there is considerable divergence as to the relation in question." (Gilbert, 2018, p. 85).

[119] As Tim Hayward (2013) has pointed out, no philosopher has yet provided a fully satisfactory account of the concept of a directed duty. However, there has been a growing trend in recent philosophical discourse to elucidate the value of the concept of a directed duty by linking it to the notion that a right-holder gains a special "standing" to object to or complain about nonperformance. This approach, as argued by Stephen Darwall (2012), emphasizes the unique position of the right-holder in holding others accountable for fulfilling their duties, as will be demonstrated in the following section. At this juncture, however, the discourse pertains exclusively to the logical possibility of directed duties; consequently, it is not contingent upon prior interaction between the parties (Darwall 2006, 21), nor does it necessitate a joint commitment entered into by them (Gilbert, 2018, p. 161).

[120] As noted in the introduction, the term "constructivist" used in this research project is deliberately broad, in order to include some peripheral authors. However, the three main authors who are unequivocally recognized as Kantian constructivists—both by others and by their own identification—are Onora O'Neill, Christine Korsgaard, and Carla Bagnoli.

than rights as fundamental" (O'Neill, 2016, p. 37). Since this is somehow a less widely accepted thesis,[121] that seems to take things "backwards", the subsequent discussion will thoroughly examine O'Neill's central argument (1.2.2) and the critiques directed towards her work (1.2.3).

Since her early writing, O'Neill proposed a Kantian constructivism wherein moral principles of justice find justification not through hypothetical consensus among rational parties, like in Rawls, but through the potential consent of real agents, maintaining a delicate balance between abstract principles and their tangible justification. "Idealized accounts of justice tend to ignore actual vulnerabilities, and relativized accounts tend to legitimate them" (O'Neill, 1989, p. 217). Specifically, in the books *Faces of Hunger: An Essay on Poverty, Justice, and Development* (1986) *Towards Justice and Virtues* (1996) and *Justice Across Boundaries: Whose Obligations?* (2016) she has directly addressed the relations of obligations and rights, underscoring the importance of taking rights seriously by emphasizing the necessity of delineating who should undertake which actions for whom (O'Neill, 1986, p. 99, 1996, p. 134, 2016, p. 10). O'Neill critiques theories of rights that fail to define clear frameworks for assigning corresponding obligations, highlighting this as a key challenge for human rights. Without justified and clearly allocated obligations to specific individuals or institutions, some welfare rights remain rhetorical (O'Neill, 2016, p. 35).[122]

[121] There are, of course, other well-known non-constructivist scholars who have expressed this view vigorously. For instance, Lawrence C. Becker argues that rights are derivative, established entirely by deriving from the (logically) precedent existence of duties in others (Becker, 1980, p. 1201). Conversely, some scholars advocate for a similarly strong but opposing perspective. Joseph Raz, for example, contends that rights can exist even without corresponding duties, particularly when they serve a justificatory role in guiding actions or shaping policies. A right to freedom of expression, for instance, may not impose a direct duty on any single individual but functions as a guiding principle for institutions and legal frameworks. Still, even Raz defines rights by using the concepts of duties, while duties are defined by referring only to the concept of reasons (Raz, 1986, p. 166). This divergence is particularly noteworthy given that both Becker and Raz are broadly categorized within the interest theory tradition. Interestingly, even within the rival framework (will theory) contrasting views on the relationship between rights and duties can be found. For example, Hans Kelsen (1967) conceptualizes a right as an individual's capacity to claim or enforce an existing legal obligation imposed by norms, effectively grounding rights in pre-existing duties, while Robert Nozick (1974) posits that rights are primary and define the boundaries of permissible actions, with duties arising secondarily to respect those rights. The nuances of these debates will be further examined in Section 1.3.1, where we delve into the distinctions and implications of will and interest theories.

[122] "If human rights were all of them liberty rights, then justice to the poor and hungry would be achieved by laissez-faire policies: provided liberties were not curtailed, justice would be served. […] But if human rights include certain social and economic rights, including 'welfare' rights, justice will require that some of these needs be met. For example, if there are rights to food or to subsistence, then it is unjust not to meet these needs, and unjust not to regulate economic activities that will prevent their being met. However, any claim to social or 'welfare' rights will be no more than rhetoric unless the counterpart obligations are both justified and allocated to individuals and institutions that can

It is important to clarify here that the critique of O'Neill to the human rights framework precedes the distinction between liberty and welfare rights and addresses a more profound one, the fact that society has nowadays focused mostly on recipient questions (entitlements) over agentic questions (requirements). While rights inherently imply obligations, these are often treated as secondary, neglecting the responsibilities necessary to fulfill rights and limiting their practical application (O'Neill, 1996, p. 127).[123] This is why it is crucial to start from obligations: "the advantage of beginning with obligations is that taking this perspective requires one to be more realistic, clear and honest about burdens, their justification and allocations[…] The symmetry of rights and obligations is evident only in the case of universal liberty rights, considered in the abstract: if A has the liberty right to do x, then everyone will have the obligation to allow A to do x. However, even in this case, the symmetry fades when institutions of enforcement are brought into the picture: A's right to do x may then be matched not only by others' obligation to allow A to do x, but by the obligations of enforcers to ensure that A is not prevented from doing x. Institutionalization disrupts any simple symmetry of rights and obligations by dispersing obligations, and sometime rights, across a plurality of agents, officials and institutions" (O'Neill, 1996, p. 135).

By underlining this problem of accountability with the "manifesto-rights", as O'Neill also calls them (O'Neill, 1996, p. 133),[124] she is not advocating for a return to a narrower conception

---

discharge them. And here many advocates of human rights are evasive about the real demands of the rights they favour. It is a significant and non-trivial matter that there is no human obligations movement." (O'Neill, 2016, p. 35).

[123] "Much contemporary writing on ethics, and especially on justice, builds on systematic deontic structures. A lot of this work is broadly liberal in orientation and the majority of it treats the perspective of the subject or recipient as prior to that of perspective of the agent, and accordingly treats rights rather than obligations or duties as the fundamental ethical notion. It is guided by the recipient's question: 'What are we (or: what am I) entitled to?' The preoccupation with recipience may be politely masked by a third-person formulation that asks 'What are human beings, or right-holders, entitled to?' (or: 'What Human Rights are there?'), but nevertheless focuses on recipience rather than on action, on rights rather than on obligations. Of course, rights do not banish obligations. On the contrary, once rights have been vindicated, obligations will be entailed: but they will not be the primary focus of ethical consideration." (O'Neill, 1996, p. 127).

[124] Manifesto rights, for O'Neill may guide agitation, but not reasonable action: "When advocates of Human Rights proclaim universal rights to food or to work or to welfare, yet fail to show who has corresponding obligations, or where claims of rights or redress may be lodged, they hurl a weapon that may boomerang. At *best* a premature rhetoric of rights may have political point and impact. An appeal to the 'manifesto rights' of the sort promulgated in Charters and Declarations invokes and highlights ideals that *may* guide agitation, politics and legislation in a quest for institutionalized, claimable rights. The resonating ideal of Human Rights, as formerly of the Rights of Man, *may* galvanize people who once conceived of themselves as mere subjects, entitled only to petition the powers that be for relief from their miseries. They *may* come to conceive of themselves as citizens, or as citizens-to-be, who can insist

of liberty rights by relinquishing the idea of welfare rights, which is undesirable for different reasons.[125] She is trying to underline that both liberty and welfare rights need institutions for their enforcements, but liberty rights don't need institutions to be reasonably claimable, while welfare rights need the institutions to conceive the assignment of obligations and redistribution.[126]

To avoid the undesirable situation created by focusing solely on the recipient's stance,[127] O'Neill draws on Kant's framework, recovering four categories of obligations, along two dimensions: universal-special and perfect-imperfect duties[128]. Perfect duties are specific, correlate

---

that justice is violated and claim what is owed to them. But at worst a premature rhetoric of rights can inflate expectations while masking a lack of claimable entitlements" (O'Neill, 1996, p. 133).

[125] In a minimalist framework of liberty-focused human rights, everything appears just, as long as individual liberties are not infringed upon. For instance, in such a framework, if a transnational corporation abruptly ceases its operations in a poor country, so devastating the local economy, no injustice has been done. O'Neill explicitly states that certain libertarian account, which interpret justice as a matter of securing the greatest possible liberty compatible with like liberty for all, render obligations without counterpart rights impossible. "The supposed 'rights to liberty pre-emptively occupies the only possible 'space' in which obligations without rights could make their demands" (O'Neill, 1996, p. 142).

[126] "It is true that rights of both sorts need institutional structure for their enforcement, but liberty rights do not need institutional structures to be claimable and waivable. By contrast rights to goods and services can be claimed or waived only if a system of assigning agents to recipients has already been established, by which the counterpart obligations are 'distributed'." (O'Neill, 1996, p. 131).

[127] For example, that the focus on recipients' question led us to conceive justice and virtues as competing rather than complementary, the reconciliation of which is the main task of her book. "This diagnosis is supported by the fact that those earlier writers of justice who did not insist on the priority of the perspectives of rights integrated account of justice and of the virtues. Locke, Rousseau and Kant, for example, whom contemporary advocates of right-based accounts of justice regard as their intellectual ancestors, begin not with rights, but with Natural Law or duty, and discuss justice and virtue in tandem. So do naturalists such as Hume and Mill, who also do not give the perspective of rights and recipience priority" (O'Neill, 1996, p. 140).

[128] Kant introduces the distinction between perfect duties and imperfect duties in his *Groundwork of the Metaphysics of Morals* and elaborates on it further in the *Metaphysics of Morals: Doctrine of Virtue*. At its core, a perfect duty is an obligation that specify the performance of an action and that must always be fulfilled without exception, while an imperfect duty do not specify actions, and thus allows for flexibility in how and when it is fulfilled, since you could not reasonably exist in a constant state of performing that duty. In Kant writings classic examples of perfect duties include the duty not to kill and the duty not to lie: no matter the circumstances of actions, you always have to act according to these principles. Conversely, the most notable examples of imperfect duties are the duty of self-improvement (4:423) and the duty to aid others. These imperfect duties are grounded in the realities of human existence and the vulnerabilities inherent to rational beings: humans are frail and dependent on others at various stages of life—whether as infants, in times of need, or in old age. Unlike contractualist moral theories however, this interdependence does not merely result in cooperation based on self-interest or agreement: the core of morality for Kant is its universalizability, thus it is also possible to universalize these personal needs to construct moral maxims. For example, "I need help at times" becomes "Every rational creature has a duty to help other rational creatures at times." This universalization creates the moral obligation to provide aid, reflecting our recognition of similar needs in others. Thus, the duty to aid others arises: I am obligated to help others when I can, just as I depend on others for help in my times of need. The second imperfect duty, self-improvement, follows naturally from this situation. For rational beings to fulfill the duty of aiding others effectively, they must develop their skills and abilities. When I need help, I require not just assistance but competent assistance—whether from a skilled doctor, a talented mechanic, or an

directly with individual rights, and are exceptionless; imperfect duties are general, do not correspond to specific rights, and allow flexibility in their fulfillment (O'Neill, 1989, p. 224). Instead, universal duties apply to all rational agents in identical situations, irrespective of personal relationships, while special duties arise from particular relationships or specific roles that individuals have.[129] From these distinctions, O'Neill represent the duties diagrammatically as follows: 1) Universal perfect duties are held by all and owed to all, corresponding to universal rights; 2) Universal imperfect duties are held by all but owed to none, with no corresponding rights; 3) Special perfect duties are held by some and owed to specific others, corresponding to special rights; 4) Special imperfect duties are held by some and owed to none, with no corresponding rights.

An illuminating example that O'Neill offers of this fourfold distinction is provided in an article (1992),[130] where she address the question of parental obligations toward children: 1) Universal perfect duties: Parents owe all others, including their children, obligations such as not to harm them—violations of which, like child abuse, infringe universal rights; 2) Special perfect duties: Parents owe their specific children care and support, corresponding to the child's right to receive these from their parents. Neglectful parents violate these duties; 3) Universal imperfect duties: Parents owe general virtues like courtesy and concern to all, including their children, even if they do not have a correlative right to it. Rigid or cold parents, while observing both their

---

experienced teacher. Therefore, the duty to perfect oneself emerges as a way of ensuring that rational beings can meet the needs of others. This duty entails cultivating one's talents and abilities, making self-improvement not merely an individual pursuit but an imperfect duty grounded in our interdependence.

[129] It is important to clarify that special imperfect duties and supererogatory actions are distinct concepts, even though the latter is more frequently discussed in the literature. Supererogatory actions, by definition, are not duties. Instead, they represent actions that go beyond what is morally required while still generating moral value. Yet, there is a sense in which supererogation is still connected with duties, as O'Neill underlines: "Saint and heroes are admired because they fulfil quite ordinary ethical requirements, but do so with superabundance. […] In a sense the clue to the notion of supererogation lies in its etymology. Supererogatory action is indeed not required, but it is measured by that which is required: in supererogatory action the ordinary measures of duty rather than the categories of duty are exceeded." (O'Neill, 1996, pp. 207–208)

[130] O'Neill has often used the figure of the teacher or parent to explain the imperfect duties they have, even from *Construction of Reasons*. "A[n imperfect] duty of this sort may be the duty to help one's children flourish. It does not entitle your child to any specific efforts on your part, for example, that you juggle for him, take him to a ballgame, or buy him a toy train." (O'Neill, 1989, p. 191). O'Neill, then, concludes that "Those who do only what the children they interact with have a [...] right to will do less than they ought. They will fulfill their perfect but not their imperfect obligations. In particular, parents or teachers who met only their perfect obligations would fail as parents or teachers" (O'Neill, 1989, p. 191).

children's universal and special rights, fail in these moral virtues; 4) Special imperfect duties: Parents owe their children unique forms of love, attention, and support specific to the parent-child relationship but to which children have no enforceable right (O'Neill, 1992, pp. 35–42).

If we take this broader picture of O'Neill's framework in mind, we see how O'Neill does not discard claim rights in favor of liberty rights, as someone might try to accuse her. She is suggesting that we need to start with the notion of duties (and specifically perfect duties) in order to understand what types of right we are referring to and the conditions under which they can be meaningfully fulfilled. Negative duties, which are duties to refrain from interfering with someone else's actions, usually correlate with liberty rights. These rights are often about the freedom to act without obstruction from others. For example, a right to freedom of speech entails that others must not prevent an individual from expressing their thoughts and opinions, meaning there is a corresponding negative duty on others to refrain from censorship or coercive suppression of speech. This is a fundamental aspect of many liberal democratic societies, where the individual's autonomy is respected as long as it does not infringe on other liberty rights.[131]

In contrast, positive duties require action or provision of something that supports or enables another's well-being.[132] These duties are typically linked to welfare rights, which assert that individuals have the right to be provided with certain goods or services (O'Neill, 1996, p. 131).[133] For instance, the right to health care imposes a positive duty on society or the government to ensure

[131] O'Neill presents imperfect duties as the core counterpoint to rights-based approaches by demonstrating that these duties are inherently open-ended and therefore do not correspond to specific rights. She refers to this openness in a twofold sense: 1) The flexibility in choosing how to fulfill a fundamental imperfect obligation and 2) The discretion in deciding for whom the obligation should be performed. She explains this distinction by stating that "What is left optional by a fundamental imperfect obligation is selection not merely of a specific way of enacting the obligation but of those for whom the obligation is to be performed" (O'Neill, 1989, p. 191).

[132] By using terms such as "welfare" and "well-being," O'Neill implicitly highlights the inherent difficulty in clearly defining what constitutes a person's well-being and how to effectively promote it. This challenge is particularly evident when well-being is treated as a master value, a concern also discussed by Scanlon (1998, pp. 141–143). We will revisit the topic of value pluralism in section 2.3.3.

[133] "Defenders of welfare rights are often eager to claim a closer analogy between liberty and welfare rights. Liberty rights too, they point out, need effective institutionalization whereby ostensibly universal obligations are in fact assigned to specific agents and agencies. The most impeccable liberty rights, such as a right not to be tortured, cannot be implemented without a legal systema and institutional structures for supervising police, courts and penal institutions. […] However, universal rights to goods and service, such as welfare rights, are in fact unlike liberty rights. It is true that rights of both sorts need institutional structures for their enforcement, but liberty rights do not need institutional structures to be claimable and waivable. By contrast rights to do goods and service cam be claimed or waived only if a system of assigning agents to recipients has already been established, by which the counterpart obligations are 'distributed'." (O'Neill, 1996, p. 131).

that individuals have access to medical services when needed. Similarly, the right to education requires governments to establish and fund schools, providing resources and infrastructure so individuals can receive an education. These are positive duties because they involve the active provision of services or resources to individuals, ensuring their basic welfare.

O'Neill thus thinks that both these types of duties and rights exist, but they correlate differently. Negative duties typically correlate with liberty rights, while positive duties align with welfare rights. Is O'Neill, however, suggesting that human rights should be considered only as liberty rights? If so, she would be presenting a minimalist view of human rights, limiting them to rights that primarily focus on non-interference, while a maximalist approach would prioritize protecting individuals from state or societal overreach, without necessarily mandating that society provide certain goods or services. This might seem to conflict with other constructivist accounts, typically regarded as maximalist, such as that of Forst (1999). Are these criticisms of O'Neill's approach valid?

1.2.3 The Critique of O'Neill's Focus on Duties

The terms minimalist and maximalist in human rights theory are subject to varying definitions. A strict definition considers any account minimalist if it posits fewer rights than those enumerated in the Universal Declaration of Human Rights (Simpson, 2021). Under this criterion, O'Neill's account would indeed be minimalist, as would Rawls'(1999). However, a less rigid definition frames minimalism as a view in which all human rights are purely negative, requiring only that others refrain from violating the rights in question. From this perspective, certain rights, such as those outlined in Articles 22–29 of the Universal Declaration (e.g., rights to social security, work, leisure, an adequate standard of living, education, and participation in cultural life), would be excluded because they entail positive duties. Conversely, on a maximalist account, every human right encompasses both negative duties (to avoid infringing on rights) and positive duties (to protect and assist individuals in enjoying their rights). From this viewpoint, there is no inherent issue with recognizing social, economic, and cultural rights.[134]

[134] A comprehensive maximalist account is presented by Henry Shue (1980), who emphasizes the interplay of duties to avoid, protect, and aid, while Joshua Cohen (2004) offers a distinct minimalist approach, arguing for a narrower conception of human rights that avoids broader obligations.

Within this more modest distinction, O'Neill would likely position herself somewhere in the middle. She does not claim that all human rights must be purely negative; rather, her argument emphasizes that any rights we institutionalize must be actionable, practically implementable, and closely tied to real-world outcomes (O'Neill, 2005, p. 427)[135] in order to hold genuine moral and practical value and not rely only on the power of condemnation and blame (O'Neill, 2005, p. 439).[136] The effectiveness of rights depends thus not merely on their conceptual formulation, but on their ability to guide and enforce duties that are intelligible and adoptable by agents within existing societal structures. Therefore, O'Neill's insistence that duties precede rights does not, in her view, undermine the practical struggles of those asserting their rights: duties arise from a deliberative process rooted in the constitutive principles of reason and shaped by empirical experiences, particularly the demands others place upon us. By deliberating on these demands, we can determine which duties we owe and assess their legitimacy. When questions about the legitimacy of these demands arise, reason itself—the ultimate source of normativity for constructivists—provides the answer.

Nonetheless, the distinctions and discrepancies some critics (Besch, 2020) have sought to highlight between O'Neill and Forst are not confined to the debate between maximalist and minimalist conceptions of human rights—a debate that is primarily about the scope of application—since there are indeed some theoretical disagreements.[137] To recapitulate, while some

[135] "What is the point of having a right? More specifically what is the point of having an abstract right, unless you also have a way of securing whatever it is that you have a right to? Why should we prize natural or abstract rights if there is no way of ensuring their delivery? And if we need to secure their delivery, are not 'the farmer and the physician' not merely of greater use than abstract or natural rights, but also of greater use than positive rights to claim food or medicine? For a hungry person, positive and justiciable rights to food are to be sure better than abstract rights that are not justiciable: but those who know how to grow, harvest, store and cook food are more useful, and having the food is better still. When we are ill, positive and justiciable rights to health care are to be sure better than abstract rights that are not justiciable: but skilled doctors and nurses are more useful, and receiving their care is better still" (O'Neill, 2005, p. 427).

[136] "I do not wish to suggest that the human rights culture inevitably promotes this rancorous approach to life. But I do not think we should accept at face value the view that it is all about respect for persons and treating others as agents. Much of it is indeed about protecting the weak and vulnerable. But it is also about extending the power of states over non-state actors and human individuals, about establishing systems of control and discipline that extend into the remotest comers of life, about running people's lives for them while leaving them with the consoling pleasures of blame. As Bernard Williams puts it, blame is 'the characteristic reaction of the morality system' in which obligations and rights have become the sole ethical currency" (O'Neill, 2005, p. 439).

[137] Thomas Besch have argued that if we follow O'Neill the duty of justification cannot be both agent-transparent—fully comprehensible and actionable from the perspective of the agent—and intelligible and adoptable by the agent (Besch, 2020, p. 377). Forst indeed explicitly distance himself from one of the two touchstones that O'Neill (O'Neill, 1996, p. 212). identifies at the heart of the constructivist approach—abstraction—focusing instead

duties do not correlate with rights, for O'Neill rights always correlate with some duties (be it negative or positive); yet this correlation is not bidirectional but unidirectional, because duties logically precede rights. Therefore, it sems to follow that, according to O'Neill, Forst should argue that the constitutive principles of generality and publicity first generate a duty of justification, which only subsequently grants agents a right to be justified. Although this reasoning might appear to impose O'Neill's framework onto Forst's system—adapting his distinctive articulation and vocabulary to fit her interpretation and lose the strength of the fundamental right to justification—Forst himself would likely accept this formulation.[138] However, he would still underscore the bi-directional relationship between rights and duties, [139] even in the case of the "fundamental right to justification" (Forst, 2012, p. 21),[140] a point where his view indeed diverges from O'Neill's (2005, p. 438).[141]

---

mostly on universalizability. "The emphasis on the importance of abstraction is also the reason why I disagree with O'Neill's claim that the constructivist approach in Political Liberalism is 'internal to a bounded society'." (Forst, 2017b, p. 130n).

[138] Scholars who focus on Forst's writings on the right of justification often overlook the broader context of his philosophical system and the central role that duties play within it. Forst (1999; 2010; 2017a) repeatedly emphasizes that practical reasonableness entails both a readiness and an ability to fulfill one's duty of justification, ensuring that one's actions and decisions meet the standards of reciprocal and general acceptability. As he puts it: "practically reasonable beings […] 'stand behind' their validity claims and duties of justification, that is, they are ready and able not only to provide adequate reasons, but also to make them the foundation of their actions" (Forst, 2017a)

[139] I'm following Laura Valentini in thinking that, although Forst does not provide a technical definition of what a right is, "his repeated references to individuals' 'authority' to demand justifications from each other suggest that the Hohfeldian idea of a claim right is in fact what he has in mind" (Valentini, 2016, p. 53).

[140] "According to the principle of reciprocal and general justification, moral persons have a fundamental right to justification, and a corresponding unconditional duty to justify morally relevant actions. This right accords to each moral person a veto right against actions or norms that are not morally justified. Each person can assert this right and demand appropriate reasons, and each person has the duty to provide them in moral contexts. The basic form of moral respect consists in observing this fundamental right; in Kantian terms, respect for moral persons as 'ends in themselves' means that one recognizes their right to justification and the duty to be able to give them appropriate reasons." (Forst, 2012, p. 21).

[141] "First-order obligation-bearers are also right-holders, and it may be that the burdens their obligations impose are recompensed by the rights they enjoy as a result of others discharging their obligations. However, the experience of right-holders is not symmetric with that of obligation-bearers. Individuals act as right-holders only when something has gone awry. In that situation they may complain, seek redress and compensation. The legislation and regulation of states that take human rights seriously often provide a range of remedies- for those with the time, energy, courage (or foolhardiness) to pursue them." (O'Neill, 2005, p. 438). For a more recent analysis of the Kantian rights-based perspective on the debate over corrective duties—specifically, the idea that corrective actions are triggered by a failure to respect others' right to freedom—see Giulio Fornaroli (2024).

Aside from these minor discrepancies, O'Neill's emphasis on the primacy of duties has faced direct criticism from scholars outside the constructivist framework.[142] Some of this criticism specifically challenges the argument that leads to the logical primacy of duty. Joel Feinberg, for example, emphasizes that rights—by empowering individuals to make claims against others—are the core concept: this power to make claims is what establishes the duties of others (Feinberg, 1970, p. 148). To illustrate the primary unidirectionality of this relationship, Feinberg asks us to imagine a world where duties exist, but rights do not. In this hypothetical world, which he calls "Nowheresville," people may be just as virtuous and as mindful of their duties, yet they lack any conception of rights. If rights were merely the flip side of duties, a world that fully understands and respects duties would be indistinguishable from one with rights. However, Feinberg argues that our world differs significantly from Nowheresville because, unlike its inhabitants, we recognize that we have the authority to demand what we are morally/legally entitled to.[143] The logical conceivability of Nowheresville is intended to demonstrate that the existence of a duty does not necessarily imply that someone else holds a corresponding right. Moreover, the key difference between our world and Nowheresville lies in the act of claiming. In our world, a legal right exists when a claim is validated by legal rules, and a moral right exists when a claim is recognized as valid by moral principles. Thus, rights are not merely the corollary of duties but are rooted in the recognition and validation of claims.

Other critics have questioned the application of O'Neill's proposal, arguing that prioritizing duties over rights may obscure the normative importance of rights as independent claims, particularly in situations where duties are unclear or contested. One of the main direct responses

[142] This critique also extends, at least partially, to scholars within the broader constructivist tradition—if we consider Henry Richardson's work as part of this intellectual landscape. Richardson (2018, 103n) argues against O'Neill by insisting that rights are co-created through active moral deliberation about what individuals are owed. In this bidirectional relationship, rights and duties are interdependent, working together to shape a just society through the practical agreement of the moral community. We will explore Richardson's proposal in more detail in section 1.3.3.

[143] Of course, a Nowheresvillian could pressure, exhort, or plead with others that they fulfill their various duties, but she would not be able to claim what is her due. As Feinberg puts it: "The most conspicuous difference, I think, between the Nowheresvillians and ourselves has something to do with the activity of claiming. Nowheresvillians, even when they are discriminated against invidiously, or left without the things they need, or otherwise badly treated, do not think to leap to their feet and make righteous demands against one another though they may not hesitate to resort to force and trickery to get what they want. They have no notion of rights, so they do not have a notion of what is their due; hence they do not claim before they take" (Feinberg, 1970, p. 148)

to O'Neill's focus on duties comes from within the Kantian framework, in Thomas Pogge's paper titled *O'Neill on Rights and Duties* (1999). Pogge is grateful to O'Neill for the revival of the duty-right distinction, but he still think that the rights idiom and the duties idiom have essentially equal potential (Pogge, 1992, pp. 237–238)[144], and he argues that it is crucial to maintain the bidirectional relationship between rights and duties, especially in the real world, where global injustices violate rights and in turn impose corresponding duties for those who benefit from or perpetuate these injustices.

He start this paper by appropriately distinguishing between the pair of claims and obligations, and the pair of rights and duties, identifying the former as contingent and the latter as unconditional (Pogge, 1992, p. 234).[145] This distinction permits to reconstruct O'Neill main argument as follow: if we aim to provide a critically reconstructive explanation of the core elements of our morality and seek to make this explanation as unified, coherent, and elegant as possible, we should base it on an account of general and universal duties. Once O'Neill's thesis is understood in this way, what initially appears to be the most powerful objection to it dissolves immediately: "In giving a central role to the notion of duties, O'Neill is not saying that what gives meaning or value to the universe is that persons do their duty, or do what they sincerely believe to be their duty. Nor is she saying that a fully moral agent cares only about trying her best, so that her duty shall have been done, and does not care about the outcome of her moral endeavors. The thesis here before us is perfectly compatible with denying these views" (Pogge, 1992, p. 236).

Pogge provides a more precise reconstruction of O'Neill's core argument than many other critics, namely that a duty-based approach captures a crucial aspect that rights-based frameworks

[144] "It is a good thing that those who - in the trenches, as it were - work for moral progress be aware of both these moral idioms, and that they feel free to use whichever idiom is most effective in communicating a particular point in some particular context. This is a suggestion we can responsibly make to them. I do not think, however, that it is either philosophically warranted or politically wise to suggest to them more than this: that by working in terms of the rights idiom they have been barking up the wrong tree all this time." (Pogge, 1992, p. 238).

[145] "A duty to keep one's promises generates obligations whenever a promise is made. Those who make no promises have no obligations. But they still have the duty: to keep any promises they make. Similarly, a right to the fruits of one's labor may generate claims to particular products of labor. Those who don't work have no claims. But they still have the right: They can acquire claims, while horses cannot. Duties and rights are fundamental vis-a-vis obligations and claims in that the former are had unconditionally while the latter are generated in particular circumstances specified by the former. Obligations and claims merely spell out what the underlying rights and duties entail under given conditions. Having a particular obligation is being duty-bound; and to violate someone's right to the fruits of her labor one must violate a particular claim of hers to some particular product." (Pogge, 1992, p. 234).

often overlook: a significant aspect of morality that can be described in terms of imperfect duties (Pogge, 1992, p. 237).[146] While Pogge does not object to the concept of imperfect duties itself, he remains unconvinced about recentering morality and politics around it.[147] Through the analysis of many O'Neill examples, Pogge detects, unpacks and ultimately rejects O'Neill's tacit assumption that rights must be specific, whereas duties can remain unspecific. Regarding the example of parental duties toward children, he argues that children could reasonably be said to have a general right to care and support for their flourishing (Pogge, 1992, p. 239).[148] Such a right, though unspecific, is no less valid than commonly recognized rights like due process or privacy, which are similarly unspecific yet widely accepted.[149]

Pogge then critiques O'Neill's example concerning world hunger and her claim that such situations are better suited to a duty-based framework than a rights-based one. He agrees that affluent individuals have a general duty to help those in extreme poverty, but no specific destitute person can claim help from a particular affluent individual. However, Pogge argues—drawing on

---

[146] "All rights entail correlative duties, but only some duties - so-called perfect duties towards others - entail correlative rights. We also have certain imperfect duties towards others which, as it happens, are of great importance in our world. They show that rights and duties are not always correlative. They also show that a duty-based presentation of morality is to be preferred, because it can include everything a rights-based presentation can include and more. The problem with the rights idiom is that in it an important part of morality describable in terms of imperfect duties - is systematically overlooked. I have sketched this argument so as to make clear that O'Neill is not accusing rights-based presentations of pretending to deliver the entire content of morality. Nor is she asserting that a duty-based presentation can deliver the entire content of morality. She may believe this. But she asserts merely that her duty-based presentation can capture something quite important that no rights-based presentation can capture." (Pogge, 1992, p. 237).

[147] Pogge underlined how O'Neill could have made the same argument simply referring to duties to oneself or genuine duties towards animals (genuine: not reducible to duties toward present or future human beings); if such duties existed, they would be obvious examples of duties without correlative rights, assuming there are no rights against oneself or animal rights. Pogge conjecture is that O'Neill does not believe there are genuine duties to oneself or towards animals, or at least "her method for constructing duties does not seem to allow their derivation" (Pogge, 1992, p. 237).

[148] "There are unspecific duties that cannot be correlated with specific rights. That was to be expected. But why not look for an unspecific correlative right? Why not say, in the case at hand, that children have a right that their parents help them flourish? There is a crucial implicit premise in O'Neill's argument, namely that rights must be specific, while duties need not be. I see no reason to accept this premise. And I see reasons not to. Many commonly recognized rights are highly unspecific: We speak of rights to due process, privacy, equal treatment, political participation, and freedom of expression. So why cannot children have a right to be cared for, or a right that their parents help them flourish?" (Pogge, 1992, p. 239).

[149] Pogge agrees that "the language of rights (and claims) might be unsuitable" in the family context, fostering an adversarial tone and exaggerating children's capacity to raise moral issues on their own behalf. (Pogge, 1992, p. 239). This is the reason why Pogge keeps repeating that he is not advocating a rights-based framework, but only that the fundamental principles of morality can be articulated equally well in the language of either rights or duties. The choice between these frameworks should depend on which is more effective in a given context.

the terminological distinction he established earlier—that this does not demonstrate the absence of correlative rights, only the absence of specific claims (Pogge, 1992, p. 240).[150] He uses an example, akin to the fictional *Nowheresville*, where all affluent people fulfill their duties, yet some destitute individuals remain unassisted. In such scenarios, those who receive help have no specific claims, and those left unhelped have no unfulfilled claims either. Nonetheless, Pogge questions whether this necessarily rules out the existence of general rights corresponding to these duties, even in the absence of specific claims. He frames his argument as follows: "Does this show that there are no rights? By no means. For O'Neill's duty can readily be translated into the rights idiom: Any person, B, has a right to be helped by any other person, A, just in case B is destitute and A is affluent and A is not already doing enough for other destitute person(s). […] This right, just like O'Neill's duty, belongs to every human being unconditionally. Only its content is conditional, just like the content of O'Neill 's duty (Pogge, 1992, p. 241).[151]

Pogge thus believes that the example of world hunger does not demonstrate the existence of duties without correlative rights, as O'Neill suggests, but rather reinforces the mutual dependence of rights and duties, that world hunger is not primarily caused by the failure of the affluent to recognize their duty to help, but by economic institutions that create a skewed global distribution of wealth and income. While more charitable actions from the affluent would reduce hunger, the core issue lies in a global economic system that predictably leaves millions in poverty and powerlessness without large-scale relief efforts. Pogge suggests that the solution is to reform

[150] "No one has a claim to be helped by you. Moreover, since the same may well be true of the other persons in a position to help, it is possible that none of the hungry has a claim against any of the affluent. Though each of the more affluent has an obligation to help, no one among the hungry has a claim that he be helped, let alone a claim to be helped by anyone in particular. Therefore, O'Neill concludes, moral situations of this kind fit comfortably into a duty-based but not into a rights-based account. This argument is flawed, because it yields a conclusion on the non fundamental level, i.e., a conclusion about claims. It concludes - correctly, I believe - that there are obligations without correlative claims, where it should be arguing that there are duties without correlative rights" (Pogge, 1992, p. 240).

[151] Pogge is not claiming that every affluent person must directly feed every hungry person: each affluent individual has obligations, but their precise scope is conditional—depending on whether they are already fulfilling sufficient duties toward others in need. This conditionality is precisely why Pogge emphasizes the role of institutions, a point on which he aligns with O'Neill (1996, pp. 129–135).

global economic institutions, framing this goal not just as a duty, but as a human right to a system where everyone can meet their basic nutritional needs (Pogge, 1992, pp. 245–246).[152]

Therefore, while both Pogge and O'Neill reject the view that aiding the hungry is optional or supererogatory (since it is a moral obligation), they diverge in how they frame this obligation: O'Neill emphasizes a duty-based formulation, while Pogge argues that it can equally be expressed as the hungry having a right to assistance. Pogge thus believes that O'Neill's position on world hunger can be restated in the language of rights without losing its moral force. (Pogge, 1992, p. 244).[153] Both frameworks—duties and rights—lead to the same practical appeals and obligations: affluent individuals are morally obligated to help the destitute, and neglecting this help violates either a right or a duty. The choice between the two depends on what is more effective in motivating action against world hunger, a judgment better made by those directly engaged in addressing the issue. In some contexts, prioritizing fundamental rights, such as human rights, which are not merely individual protections in isolation but represent entitlements embedded within broader social institutions and practices, may be particularly apt (Pogge, 1992, pp. 246–247).[154]

---

[152] "[W]e should not see the problem of world hunger as primarily due to the fact that too few among the affluent recognize and live up to their duty to help those too poor and impotent to help themselves. Rather, we should see it as due primarily to economic institutions that engender a reasonably stable and rather skewed global pattern of income and wealth distribution. To be sure, there would be less malnutrition and starvation if more of the affluent extended a helping hand. Nevertheless, there is something deeply unjust about a global economic regime whose centrifugal tendencies ensure, quite predictably, that without massive relief efforts millions will end up extremely poor and powerless. Now if this explanation […] is essentially correct, then our chief goal must be to change our global economic institutions. This goal can, of course, be described as a duty. But it can also be described in terms of a human right […] to a global economic regime under which all can meet their basic nutritional needs. Realizing this right probably does not require the creation of a legal right to food-when-needed. But it does require some modifications in our global market system - perhaps a reconceiving of property rights so that every human being is seen as entitled to some share of our planetary resources, a share that could be underwritten through charges on resource extraction, pollution, and the like." (Pogge, 1992, pp. 245–246).

[153] "The rights idiom allows us to make the same appeal in different words: "You are affluent and others are destitute. If you help none of them, then you are violating a right that each of them has against you. By helping some of them, you can avoid violating the rights of any"." (Pogge, 1992, p. 241).

[154] "By proposing this understanding of our most fundamental rights, I am not denying Bentham's point (which O'Neill quotes prominently […]) that "a declaration of rights would be but a lop-sided job without a declaration of duties". Of course it would. Correlative to human rights there are, at the same most fundamental level, universal duties: Those who participate in the imposition of unjust social arrangements must help to improve such arrangements and help to prevent and mitigate their worst effects. These human rights and duties may be conceived as coeval: two sides of the same coin, neither more basic than, or grounded in, the other. My point is not to support the primacy of the rights idiom, but merely to rebut O'Neill's arguments for primacy of the duty idiom. This point, I have argued, is

Many constructivists align with Pogge's critique of O'Neill's emphasis on the primacy of duties over rights. While constructivists acknowledge that duties and rights do not always correlate (as in cases of imperfect duties), they reject the notion that duties inherently take precedence. Instead, when rights and duties do correlate, they are bidirectional and interdependent, forming a dyadic relationship rather than a hierarchical causal relationship (Richardson, 2018, p. 23).[155] In the following section, we will examine how a constructivist framework that embraces this reciprocity between rights and duties provides a more coherent foundation for addressing both theoretical and practical moral and political challenges.

1.3 THE FUNCTION: SPECIFICATORY DEMAND

1.3.1: Will and Interest Theory

Even if we agree that claim-rights are the predominant form of rights, the debate becomes particularly lively when considering their functions,[156] leading to the emergence of two main distinct currents: will (or liberty or choice) theorists and interest (or benefit or desire) theorists.[157]

---

supported by the fact that the problem of world hunger - whether we analyze it her way or my way does not bear out her assertion that there are duties without correlative rights." (Pogge, 1992, pp. 246–247).

[155] As we will see in paragraph 1.3.3, and specifically footnote n.206, Richardson identifies at least three reasons for favoring a symmetrical specification of rights and duties between two agents whenever feasible (Richardson, 2018, p. 112).

[156] These functions are usually expressed in at least a threefold sense: purpose, protections and specification "Each rights theory, Interest and Will, contains three features: (I) an ultimate purpose that all rights (or at least those covered by the particular version of the theory) purportedly serve; (II) means of protecting or effectuating that purpose, e.g., by imposing duties on other parties, by including enforcement mechanisms, etc.; and (III) specifying which candidates constitute rights, i.e., which types of normative positions (e.g., Hohfeldian claims, liberties, etc.) and which tokens of those types (e.g., what subset of Hohfeldian claims) count as rights" (Frydrych, 2019, p. 456). I will supplement Frydrych's argument with Richardson's (2018, p. 28) in order to argue that, from a constructivist perspective, the primary (though not the sole) function of a right is specification.

[157] To be more specific, even if they are often used interchangeably, choice theories are a subset of will theories (every choice theory is a will theory, but not every will theory is a choice theory) that emphasize that rights give individuals the power to make choices, particularly the power to enforce or waive the duties of others. Similarly, Desire Theory can be viewed as a specific type of Interest Theory where the primary focus is on desires (what individuals want) while Interest Theory encompasses a broader range of factors that might affect an individual's well-being or rights, including but not limited to desires. In essence, choice theories focus on the specific aspect of will theories that involves the right-holder's capacity to make choices or decisions about how the right is exercised, while

Will theorists argue that rights primarily serve to allocate domains of freedom, emphasizing personal autonomy and decision-making, giving the right-holder a measure of control over the duty of another. In contrast, interest theorists assert that rights primarily function is to protect an interests of sufficient importance to impose on other duties, serving as passive safeguards "to further their holders' interests" (Wenar, 2005, p. 240). [158] Put concisely, we can say that for interest theory the main function of a right is the content of interest it protects, while for the will theory the main function of a right is the shape of will it express. [159] This long-standing division, deeply rooted in legal theory, is exemplified by prominent proponents such as Hart (1961), Wellman (1985), Sumner (1987), and Steiner (1994) for will theory,[160] and Austin (1832), MacCormick (1982), Raz (1984), and Kramer (2013) for interest theory.[161]

---

desire theory, by emphasizing desires, adds precision to the broader concept of interests used by interest theory, suggesting that what individuals want is crucial for determining their well-being.

[158] Some scholars have traditionally framed the debate between Will Theory and Interest Theory as a matter of form. Will theorists, who view rights as providing the right-holder with choices or deliberative power, typically conceive of rights as pairs of privileges and powers. In contrast, proponents of Interest Theory argue that rights exist primarily to protect and further the right-holder's interests, framing rights as essentially claims and immunities. However, this characterization has been challenged by Leif Wenar's work, which distinguished between analyzing the form of rights (as defined by Hohfeldian incidents) and their function for the first time (Wenar, 2005). According to Wenar, while legal or moral rights can take the form of privileges, claims, powers, immunities, or combinations thereof, these forms serve one or more of six distinct functions: exemption, discretion, authorization, protection, provision, and performance. However, he further contends that neither Will Theory nor Interest Theory sufficiently accounts for all six functions of rights (Wenar, 2005, p. 246), thus calling for a more nuanced understanding in his Several Functions Theory. "All rights are Hohfeldian incidents. All Hohfeldian incidents are rights so long as they mark exemption, or discretion, or authorization, or entitle their holders to protection, provision, or performance. Therefore, rights are all those Hohfeldian incidents that perform these several function" (Wenar, 2005, p. 252).

[159] For a comprehensive discussion on the will and interest theories, refer to the work of Steiner, Kramer, and Simmonds (2002). For an argument specifically addressing the circularity of this debate, see Cruft (2004).

[160] It is noteworthy, for our broader analysis of constructivism and Kantian constructivism, that Kant is often considered one of Will Theory's foundational figures (Steiner et al., 2002, pp. 135–136). However, Arthur Ripstein (2009) has argued that the instrumentalism underlying Will Theory (and interest theory) is fundamentally incompatible with core Kantian principles, since it undermines the inherent authority and stability of rights by treating them as contingent tools for achieving external goals rather than as intrinsic to the legal and moral order. In this Kantian reading, rights and their correlatives are not malleable instruments for broader objectives but integral components to the autonomy of individuals within the legal system. As he puts it: "Kant's conception of the right to independence rests on neither of what is referred to in recent literature as [the] "interest theory" or "will theory" of rights. Underlying the other differences between these accounts is a shared conception of rights as institutional instruments that constrain the conduct of others in order to protect things that matter apart from them. […] If rights are understood in this instrumental way, they are always at least potentially conditional on their ability to secure the underlying values that they are supposed to protect. The Kantian right to independence, by contrast, is always an entitlement within a system of reciprocal limits on freedom, and so can only be violated by the conduct of others, and its only point is to prohibit that conduct." (Ripstein, 2009, pp. 34–35). As will be restate later, I believe that the majority of Kantian constructivists align with Ripstein in rejecting the premises of Will Theory and in framing the debate solely in terms of Will/Interest.

We saw how Hart (1955) criticized Austin's conception of rights as derivative from sovereign commands, arguing instead for a more nuanced understanding of rights that emphasized individual autonomy and agency: this perspective established the first core of what is now considered the Will Theory. For will theorists, that followed Hart, the essence of a right lies in the ability to make choices about how that right is exercised or waived, emphasizing the right-holder autonomy and power to enforce or modify corresponding duties. This capacity to make choices was central: a right was not merely a protection but a domain of agency in which the rights-holder could determine the extent and manner of their entitlement. For instance, if a person A has a right to property, A can choose to sell, retain, or gift that property, and all these choices define the meaningful exercise of the right.

Joseph Raz, in *The Morality of Freedom*, contested Hart conception and provided an alternative definition of the function of rights that runs approximately as follows: X has a right against Y if and only if X is capable of holding rights, and, all else being equal, an aspect of X's well-being (his interest) is a sufficient reason for holding Y to be under a duty (Raz, 1986, p. 166). In this framework, the justification for imposing duties on others stems from the protection or promotion of significant aspects of the rights-holder's well-being, rather than their autonomy. Rights are instruments for safeguarding important interests, such as health, security, or freedom from harm, and these interests provide a moral basis for the duties of others. For example, children have a right to education because education significantly contributes to their well-being, even though they may lack the autonomy to demand or waive this right. The focus shifts from the rights-holder's capacity to exercise control (as in will theory) to the importance of their interests in justifying duties (Raz, 1986, p. 247).[162]

[161] Joseph Raz frequently references his interest theory and Wellman's will theory as paradigmatic examples within the broader debate on the nature of rights (Raz, 1986, pp. 165–192). However, many scholars now regard Matthew H. Kramer's version of the Interest Theory (Kramer, 2013), which outlines specific conditions necessary for holding a legal right, as the leading contemporary alternative (Wenar, 2023).

[162] Raz's argument does not challenge the claim that the importance of an individual's interest in living an autonomous life can ground rights with correlative duties to provide standard protections of this interest, but he believed that an individual's interest in autonomy, no matter how great, could not justify holding "members of the society at large to be duty-bound to him to provide him with the social environment necessary to give him a chance to have an autonomous life.

Carl Wellman sought to refine Hart's Will Theory by addressing the limitations highlighted by Raz while at the same time preserving what he regarded as its core principle: that rights are primarily designed to enhance individual autonomy. In Wellman's view, all real rights are not static entities but dynamic complexes composed of varying combinations of Hohfeldian positions around a central core (Wellman, 1985, p. 59, 1995, p. 100, 2011b, p. 18).[163] They are dynamic because this core can vary over time (Wellman, 1995, pp. 7–9, 1995, pp. 190–195), but despite their dynamism every right contains at least one liberty and one power (Wellman, 1985, pp. 95–96, 204–211, 1995, pp. 108–112, 186–187). This reading underscores the active and empowering function of rights, to exercise control over obligations - through enforcement, waiver, or modification - which distinguishes Will Theory from Interest-based accounts. However, Wellman acknowledged certain exceptions to this framework, such as the rights of children, individuals with severe cognitive impairments, or those otherwise unable to exercise control over their entitlements. While such exceptions posed challenges to the universality of the autonomy-based model, Wellman argued that they should not fundamentally redefine the theory; instead, he treated these cases as special circumstances that require supplementary explanations but do not undermine the general structure of rights which lies in protecting and affirming the right-holder's identity or status, not necessarily the interests underlying the right (Wellman, 1995, p. 124, 2011b, p. 22).[164]

Many contemporary scholars have significantly advanced this debate (Kramer & Steiner, 2007), yet a growing number have questioned whether framing the discussion in this way remains meaningful or productive (Tasioulas, 2010, p. 656),[165] and whether the debate itself is founded on

[163] Singular Hohfeldian positions (e.g., a claim, a privilege, etc.,) are thus never rights in or by themselves (Wellman, 1985, pp. 59–60).

[164] Wellman presents Will Theory as a gradualist account of moral obligations, illustrating this with the thought-provoking "Suckling Infant Case". As he writes: "[It] would be idle and misleading to ascribes moral rights to any being incapable of action. Therefore, a necessary condition for the possession of any moral human right is the capacity for moral action of any normal adult human being. But some biological human beings, especially neonates and the irreversible comatose, are incapable of acting, in any morally relevant sense. Therefore, it is only beings that possess the normal human capacities for acting who could possibly possess any moral human right. It may seem monstrous to deny that infants can have any moral human rights. But no moral philosophers, and certainly no parent, imagines that a sucking babe has a moral duty not to inflict pain upon its mother by biting her breast or that a very young child has a moral obligation to refrain from throwing disliked food onto the floor. We all realize that children acquire moral obligations gradually as they grow to become responsible moral agents. Why not recognize that similarly infants acquire their moral rights only as they develop their capacities?" (Wellman, 2011b, p. 22).

[165] Another way to frame the debate (Wenar, 2023) is by distinguishing between two primary approaches to rights, known as 'Duty-Based' or 'Status' theories and 'Goal-Based' or 'Instrument' theories, which parallel the classical distinction between deontologist and consequentialist (Dworkin, 1977). Duty-Based theories operate on the

a flawed assumption—that there exists a single, unified concept of a right for which these theories offer competing analyses (Hayward, 2013). Leif Wenar is probably the most notable critic of both theories for oversimplifying the complexity of rights, which he sees as systematically ambiguous and context-dependent (Wenar, 2005).[166] He believes that the features of rights captured by the theory of will and the theory of interest are complementary rather than mutually exclusive (Wenar, 2005, p. 244)[167]; thus, he proposes an alternative pluralist approach, the *Several Functions Theory*,

premise that rights are fundamentally grounded in or derived from a person's status which, although it might be interpreted in various ways, it is most commonly understood as human dignity. Because human dignity is inherently unwaivable and inalienable, no one is ever justified in violating another person's rights. For example, Robert Nozick (1974) argues that rights function as 'side constraints' on the pursuit of good outcomes. These side constraints, rooted in a deontological understanding of human dignity, serve to protect individuals from the so-called 'utilitarianism of rights'—the idea that there might be circumstances, such as preventing a catastrophic harm, where infringing on someone's rights could be justified. Goal-Based theories, in contrast, view rights as tools for achieving an optimal distribution (e.g., maximization, equality) of certain advantages (e.g., utility, capabilities, resources). Proponents of this approach argue that there may be situations where infringing on someone's rights is not only justified but necessary to achieve a greater good. Even some scholars that are critique of utilitarianism like Judith Thomson (1990) introduces the concept of threshold rights, suggesting that one is justified in overriding a right if and only if the infringement produces a significant amount of good. In this view, maximizing good outcomes—such as sacrificing one innocent person to save five others—is a widely accepted criterion for determining when it is permissible to override someone's rights (e.g., the rights of the innocent individual). The line between these two approaches is often blurred because many philosophers mix their micro-level accounts of rights (interest-will) with macro-level explanations of their functions (deontology-consequentialism). For example, philosophers may share the same micro-level perspective while defending different macro-level accounts: early utilitarians, like Bentham, saw rights as instruments for maximizing societal utility (a Goal-Based Theory) and inferred from this that rights benefit their holders (the Interest Theory). However, more recent Interest theorists (Kramer, 2000; N. MacCormick, 1982; Raz, 1986, p. 198) argue that rights serve to protect and enhance individuals' well-being, which they often interpret through the deontological lens of human dignity (aligning more closely with a Duty-Based Theory).

[166] Wenar explains that the diversity of functions associated with rights arises from their historical evolution. Over nearly two thousand years, the concept of rights has been reshaped by strong social forces, from ancient law to medieval Christian debates, to the modern bureaucratic state, and by popular movements fighting for protection and empowerment. This long and varied history has resulted in an irreducibly complex concept of rights. Attempts to simplify or reduce the concept to a single logical form or function, Wenar argues, have consistently failed. "The standoff between rights theorists in the twentieth century showed that attempts to present some part of the concept of rights as though it were the whole concept will not succeed. Any analysis that cannot imagine the rights of judges and traffic wardens, of the comatose and of children, will not survive outside of the study. Any theory that must retreat to a technical language of "strictly speaking" is not, strictly speaking, a theory of rights. Theorists who try to tie reasoning about rights to a single pillar will bind their readers with ropes of sand." (Wenar, 2005, p. 252).

[167] "The will theory and the interest theory are both inadequate to our understanding of rights, the weakness of each being the strength of the other. The will theory captures rights that give discretion to the rightholder without conferring benefits, but fails to capture rights that confer bene- fits without giving discretion. The interest theory accepts rights that confer benefits, but rejects rights whose holders do not benefit from holding them. […] Each theory includes rights that the other cannot." (Wenar, 2005, p. 244). Wellman criticizes Wenar's critiques of Hart and MacCormick for failing to distinguish between essential and non-essential functions of rights. Wellman argued that Hart already distinguishes, although not explicitly, between "the essential or definitive function of rights", recognizing a choice of the right-holder, and "the non-essential or conceptually contingent functions" of conferring a liberty, immunity, claim or power upon her (Wellman, 2011a, p. 170).

that seeks to reconcile the inclusiveness of the interest theory with the emphasis on agency in the will theory, by distinguishing six functions: exemption, discretion, authorization, protection, provision, and performance. Rights are all those Hohfeldian incidents that perform these several function (Wenar, 2005, pp. 235, 246).[168] "The several functions theory thus resolves the standoff between the will theory and the interest theory by accepting all of the rights that are recognized by each of the procrustean single-function theories." (Wenar, 2005, p. 248).[169]

Furthermore, regarding specifically claim-rights, Wenar argues that the traditional "role theory", which ties rights to specific social roles, is no longer adequate in modern contexts where rights are recognized for humans, children, and even animals—entities that do not necessarily fit into predefined societal roles.[170] To address this limitation, Wenar proposes an expanded framework centered on the term "kind" (derived from the Latin *genus*), which denotes a set of entities sharing defining characteristics. This concept is broader than "role", encompassing both natural kinds and social kinds; for example, parents are members of both a natural kind (as human beings) and a social kind (as parents). The Kind-Desire Theory is an application of the insights from the Several Functions Theory (which, to repeat, is a general framework for understanding rights' various roles across contexts), for defining and grounding claim-rights within a system of norms in which some entities are conceptualized as kind. Within such a system, claim-rights

[168] "The several functions theory holds that any incident or combination of incidents is a right, but only if it performs one or more of the six functions […] In other words, all (combinations of) incidents are rights so long as they mark exemption, discretion, or authorization, or entitle their holders to protection, provision, or performance. Only those (combinations of) incidents that perform at least one of the six specific functions are right" (Wenar, 2005, p. 246)

[169] "The several functions theory accepts that there is no one thing that rights do for rightholders. Rights have no fundamental normative purpose in this sense. Rather, rights play a number of different roles in our lives. Some rights give the rightholder discretion over others' duties, some rights protect the rightholder from harm, some rights do neither of these things but something else altogether. All rights perform some function, but there is no one function that all rights have. There is no one function that all rights have, as there is no one function that all furniture has, and no one feature that all games have. Rights have all of the functions that they do" (Wenar, 2005, p. 248).

[170] Wenar follow Arthur Applbaum critiques of the role-based perspective, suggesting it evokes outdated notions of hierarchical social orders, making such talk "suspect to liberals who have no nostalgia for societies divided into castes, estates, or classes, and who fear retreat from the idea of universal humanity" (Applbaum, 2001, p. 47)

correspond to those enforceable strict duties that the members of the relevant kind want to be fulfilled, underlying the relational dynamics of kinds (Wenar, 2013, p. 219).[171]

The Kind-Desire Theory presents an innovative approach to grounding claim-rights by linking them to the desires of specific kinds within a system of norms. However, it faces significant critiques and potential weaknesses (Wellman, 2011a), and its compatibility with the broader constructivist framework, which is what we are investigating, seems to be difficult (Richardson, 2018, pp. 91–92).[172] Nevertheless, Wenar's theory has undoubtedly opened the door to other promising alternatives that challenge the conventional distinction between will and interest theories. To fully appreciate the innovation of these 'third theories', it is useful to briefly revisit again the key differences of Will and Interest Theory, particularly in the context of contractual obligations (where these theories where firstly codified) to reenter in the discussion of the dyadic relationship we are interested in. This will highlight how alternative theories can address the limitations of both approaches without merely reverting to one or the other.

Will theory, for instance, emphasizes the significance of the parties' intentions or choices in contractual agreements, arguing that a contract is binding because the parties have voluntarily consented to its terms. The focus is on the subjective intent of the parties at the time of contract formation; they are obligated to fulfill their promises because they have freely chosen to do so,

[171] Wenar formalize the kind-desire theory as follows: "Consider a system of norms S that refers to entities under descriptions that are kinds [or roles], D and R. If and only if, in circumstances C, a norm of S supports statements of the form: 1) Some D (qua D) has a duty to phi some R (qua R); where "phi" is a verb phrase specifying an action, such as "pay benefits to," "refrain from touching," "shoot," and so on. 2) Rs (qua Rs) want such duties to be fulfilled; and 3) Enforcement of this duty is appropriate, ceteris paribus; then: the R has a claim-right in S that the D fulfil this duty in circumstances C.' (Wenar, 2013, p. 219). Therefore, in order to have a claim right we need these three conditions: 1) the existence of a duty, where one kind D has duty toward a kind R (for example the kind parents have a duty to provide food to the kind children); 2) the desire for the duty to be fulfilled, where members of the kind R want the duty to be fulfilled (the duty to provide food aligns with what the children needs); 3) the enforcement is appropriate, where it is reasonable, all else being equal, to enforce this duty (society agrees it's right to hold parents accountable for not feeding their children). When these three conditions are satisfied, the members of R (e.g., children) have a claim-right to the action (e.g., being fed) from the members of D (e.g., parents).

[172] Especially in the Kantian constructivism framework. For example, this is how Wenar speaks of Kant, which is viewed as one of the fathers of Will theory: "Both will and desire are voluntative concepts, yet, to use Kant's terms, claim-rights concern the holder's wishes not her choices. The powers of promisees and owners do make brief cameos within the kind-desire analyses of their respective rights. Yet the formalization shows that powers need not be the partners of all claims. This makes kind-desire theory immune to the standard counterexamples to the Will Theory, such as unwaivable rights, incompetent right-holders, and citizens under the criminal law." (Wenar, 2013, p. 226). Against this reading, see Ripstein (2009) and see footnote n.160.

giving them legitimate claims regarding the contract's formation.[173] The function of a right is to give the right-holder control over another party's duty (Hart, 1983). A promissory right is thus the typical exemplar of a right for the Will Theory, because it provides the promisee with an option to either waive or enforce the performance of the promised action. (Wenar, 2013, p. 202). Interest theory, on the other hand, emphasizes the objective interests or benefits of the parties involved in a contract: contracts are enforceable because they serve to protect and promote the parties' interests. This theory looks beyond the parties' subjective intentions and instead considers whether enforcing the contract would serve broader societal interests or promote fairness and efficiency. The function of a right is to legitimize an essential human interest over another party's duty (Raz, 1986).[174] The representative instance of a right for the Interest Theory are the property rights, since they protect and further the owner's interests in ownership. "Property is a right because ownership makes people's lives go better" (Wenar, 2013, p. 202).

Thus, both will theory and interest theory offer distinct explanations for the function of rights in relation to their holders, each with its own strengths and weaknesses. Will theory, for example, emphasizes the intrinsic link between rights and normative control. According to will theorists, possessing a right means having the authority to determine what others may or may not do, thus exercising control over a specific domain. This connection between rights and authority—particularly the authority to regulate others' actions—is central to the will theory and is seen as a defining characteristic of rights, which are therefore characterized mainly as *instruments of control*. In the interest theory instead, rights function primarily as *protectors of welfare*. They exist to ensure that individuals' fundamental interests are protected from harm or neglect. The role of rights, therefore, is to provide the necessary conditions for individuals to thrive. This divergence is manifested in the resolution of conflict, where will theory approach conflicts of rights by looking at whose autonomy should prevail (determining whose control or decision-making should be

[173] The centrality of contracts in defining a right stems from the widely accepted definition of a legal person used by most legal scholars—namely, any entity capable of performing actions typically associated with a human person in law, such as entering into contracts, suing, and being sued (B. Smith, 1928, pp. 283–25). The focus on contracts is also central to certain conceptions of moral person, as seen in the work of Scanlon (Scanlon, 1982, 1998). However, in this context, the parties whose agreement is in question are assumed not merely to be seeking some kind of advantage but also to be motivated by the aim of finding principles that others, similarly motivated, could not reasonably reject (Scanlon, 1998, p. 5).

[174] We have a right if and only if we have an interest "that is, other things being equal, sufficient to justify holding some other person to be under a duty." (Raz, 1986, p. 32).

prioritized) while interest theory resolves conflicts by considering which interests are more fundamental or pressing (determining which rights protect the more significant or urgent interests and prioritizes those in the resolution of conflicts)[175].

Another one of the main significant divergences between the two theories is the debate over alienability. Will theorists argue that rights are, like in any contracts, inherently alienable, meaning they can be transferred or waived; in contrast, interest theorists contend that alienability is not a necessary feature of rights and that some rights must be inalienable to safeguard the most crucial interests of their owner.[176] This "moral appeal" for inalienable rights is used by interest theorists to challenge the adequacy of will theory, which fails to account for the protection of fundamental interests (Kramer, 2000). Some proponents of will theory have argued, in response, that within Hohfeld's framework, inalienable rights cannot exist since that would lead to a problematic infinite regress, as inalienability is, in Hohfeld's terms, an immunity that correlates with the disabilities of others (Steiner, 2013).[177]

Both theories, thus, exhibit certain restrictive tendencies, albeit in different ways. The interest theory tends to be more restrictive and paternalistic by not allowing the waiver of certain rights and by resolving conflicts in a manner that may override the agent's will. Additionally, the interest theory struggles to account for rights that cannot be directly justified by individual interests; this limitation prevents interest theorists from adequately explaining rights associated with social roles, such as a journalist's right to protect their sources, a police officer's right to stop and search individuals, or a judge's right to sentence criminals. On the other hand, the will theory can also be restrictive, not only because it struggles to account for unwaivable rights (freedom from enslavement is not a right if I can decide to voluntarily enslave myself) but because by tying

[175] This parallel is often cited as a critique of interest theory, as balancing and prioritizing various interests can be more complex and contentious. Richardson specifically challenges Interest and Desire theories for being "mysterious" and lacking clarity (2018, p. 92).

[176] An interest theorist offering an argument in favor of inalienable rights is Neil MacCormick (1977).

[177] However, Pierfrancesco Biasetti (2015) has offered a compelling counterargument to Steiner by demonstrating that the issue of regress is inherent throughout Hohfeld's schema, not just in the case of inalienable rights. Biasetti suggests that the most reasonable solution is to relax the strict Hohfeldian interpretation of oppositions by introducing a degree of incompleteness into the system of Hohfeldian incidents. This allows for situations where, for a particular pair of individuals and a specific action, neither of the opposing rights or duties may hold, thereby avoiding the infinite regress issue altogether (Biasetti, 2015, p. 151).

rights to the capacity for autonomy and rational decision-making, those who cannot exercise will (such as infants, people with severe cognitive disabilities, or even groups of individuals) might be seen as having fewer or different kinds of rights. In contrast with this final point, the interest theory can, in principle, extend rights even to animals by focusing on the protection of interests. Yet, even the inclusiveness of the interest theory also presents challenges: by potentially recognizing rights based on a broad range of interests, there is a risk of overextending and inflating the concept of rights: not all interests are equally significant, and without careful definition, the interest theory could lead to an excessive proliferation of rights, making the framework unwieldy and complicating the prioritization and enforcement of these rights.

The seemingly interminable debate between will and interest theories has not only brought to light the inherent limitations and challenges associated with each individual approach, but it has also stimulated a growing number of scholars and theorists to seek out innovative alternatives, as we said. These emerging perspectives aim to integrate various aspects of moral, social, and legal reasoning into a more comprehensive understanding of rights, in order to offer a more nuanced frameworks that address the complexities and contradictions inherent in the debate. In the following sections, I will examine some of these alternative approaches, with a particular focus on identifying a constructivist thread that unifies them.

1.3.2 Darwall and the Demand Theory of Right

One of the most notable alternative to Interest and will theory is referred to as a “demand theory of rights”, which emphasizes that the main function of rights involves the ability to make demands (Herstein, 2023).[178] This theory reasserts the primary significance of the claim-right form and ties it closely to agency, portraying rights-holders as agents who can insist on the fulfillment of their rights by making demands on others: having a right entails, in short, the capacity to demand certain behavior from others. The key proponents of this theory are Joel Feinberg (1970, pp. 245–

[178] Some authors of demand theory, such as Margaret Gilbert, will differ from Herstein classification, by considering demand as the form rather than the function of right (Gilbert, 2018, p. 57). In her formulation, "X has a demand-right against Y if and only if X has the standing to demand of Y that Y φ" (2018, 70). Other authors do not view demand theory as an alternative to will and choice theory, but often regard it as a refinement or elaboration of will theory (Kurki, 2019).

254, 1980, pp. 130–158, 1992, p. 155)[179] and Stephen Darwall (Darwall, 1983, 2006, p. 18, 2007, pp. 60–65), and more recently John Skorupski (2010) and Maragaret Gilbert (2018).[180] In this discussion I will focus only on Darwall, as he can be viewed as a supporter of the broader constructivist proposal (Darwall, 2006, pp. 293–297).[181]

Darwall's conception of "demand-rights" is centered on the idea that rights inherently involve a unique type of interpersonal, or "second-personal," rational authority: one makes a claim on someone's will by addressing to the person in question a certain kind of reason, a second-personal reason.[182] Darwall attempts to elucidate the concept of second-personal reason by analyzing the authority relations that an addresser assumes exist between themselves and their addressee. For a second-personal reason to possess genuine moral weight, it must be grounded in *de jure* authority, meaning the addresser must have a legitimate moral standing to expect compliance. In the absence of such legitimacy, when authority arises for example from power or influence (*de facto* authority), it lacks moral legitimacy and doesn't create a binding moral reason

[179] For Feinberg, holding claims is intimately related to human dignity, as it engages people's agency and capacity to make claims and provides them with a sense of self-worth as creatures who are not only recipients of moral regard, but are also owed it (Feinberg, 1970, p. 252). We will return to the role of dignity for constructivists in Chapter 3.

[180] Skorupski argues that rights specify what the right-holder is entitled to demand from others, where "demand" implies not only the right to request compliance but also the authority to compel performance or seek compensation in cases of non-performance. This proposal underscores the enforceable nature of rights, emphasizing the legitimacy of holding others accountable when they fail to meet their obligations (Skorupski, 2010, pp. 24–32). Gilbert, on the other hand, specifically highlights how joint commitment should be considered the only source of demand-rights: "It may be that simply seeing another as having demand-rights is a matter of respecting them. In that case those who mutually recognize each other automatically respect one another. […] [M]utual recognition, as constituted by the pertinent joint commitment, gives us demand-rights to the kind of behavior that morality is often deemed, crucially, to require of us." (Gilbert, 2018, pp. 259–260).

[181] Though Darwall does not think that his view requires metaethical constructivism, he claims it is a view he would "happily" embrace (Darwall, 2007, p. 64). O'Neill, in reviewing one of his first book, *Impartial Reason*, writes that "his constructive account of practical reason is broadly Humean in some ways and Rawlsian in others" (O'Neill, 1987, p. 62). Other scholars (Samuel, 2023) have used the term "Hegelian Constructivism" or even "Fichtean Constructivism" to describe Darwall's approach.

[182] This suggests, at the very least, that the other person is capable of exercising their will based on such reasons, thereby highlighting the importance of addressing someone who can act according to their own decisions and preferences. However, the focus here is not merely on acknowledging the individual's will but on directly appealing to their capacity as an autonomous and rational agent. "Whenever someone addresses a second-personal reason of any kind, including just proposing an offered agreement […] he presupposes that the other has the second personal standing of a free and rational agent." (Darwall, 2006, p. 200). In addressing you second-personally, I indicate that I regard you as capable of determining your actions based solely on the nature of my address, namely I convey that I consider you generally free (at least in the sense of being able to determine your actions) and rational (at least in the sense of being responsive to second-personal reasons in deciding what to do).

in the same way; it's situational and limited to that particular context (Darwall, 2006, pp. 52–54).[183]

This authority, according to Darwall, is not just about expecting compliance from others, but also about having the power to protest, resist, or even require compensation when these rights are violated. For example, if someone steps on your foot, you not only have the right to expect them to avoid doing so again but also to demand they refrain from doing it in the first place, and potentially seek restitution if they fail to comply (Darwall, 2006, pp. 5–8, 128–129).[184] However, Darwall acknowledges, like other scholars, that the primary difficulty lies in the tension between demanding compliance from someone and the expectation that this demand is appropriate regardless of the other person's inclinations. In this sense, the legitimacy of reconciling directed duties with the normative constraints they impose is the main "*demand-right problem*" (Gilbert, 2018, p. 79).[185] As Darwall puts it, "someone can sensibly accept this second reason for moving his foot, one embodied in your claim or demand, only if he also accepts your authority to demand this of him (second personally). That is just what it is to accept something as a valid claim or demand" (Darwall, 2013c, p. 137).

Darwall's solution is that for a demand to be morally legitimate both parties must recognize the normative structure that underpins it. When one person issues a demand, they do so from a position of authority that is not arbitrary but is instead grounded in the recognition of the other's moral standing. In turn, the addressee must accept the authority of the demand and the legitimacy of the right-holder's claim. Unlike mere rules or contextual norms that people ought to follow, demand-rights fundamentally entail a relational structure that grants rights-holders both the power

---

[183] Darwall draws on Austin's framework of speech acts to distinguish between genuine (*de jure*) and stipulated (*de facto*) authority relations (Darwall, 2006, pp. 52–54).

[184] In this framework, Darwall explicitly builds on insights from Feinberg and Hohfeld. Feinberg describes rights as giving us a "place to stand," meaning that rights provide individuals with a distinct moral standing from which they can make claims (Feinberg, 1966). Similarly, in Hohfeld's analysis, a right-holder possesses a "claim right" against others, which means those others hold a corresponding obligation specifically *to* the right-holder. This relationship is not simply that others should refrain from certain actions, but that they *owe* it directly to the individual, creating a specific interpersonal accountability.

[185] "How are demand-rights possible? In particular, how—at the deepest level—is it possible for one person to accrue the standing to demand of another that he perform a given action? I call this the demand-right problem or, more fully, the general demand-right problem." (Gilbert, 2018, p. 79).

to assert claims and the authority to expect others to recognize these claims, thereby reinforcing their moral standing within a moral community of equal, autonomous human agents who hold each other accountable. The authority of the demand is therefore a relational authority, not an isolated request but a claim embedded within a broader set of moral relationships that are acknowledged and respected by all members of the community.

It is important to underline that the legitimacy of this authority is not rooted in the social sphere in a weak sense but is logically grounded in reciprocal recognition and the universal moral equality of individuals. Each person possesses the capacity to make claims on others and demand moral consideration: however, for one individual to legitimately issue a demand to another, they must first acknowledge the addressee's equal moral standing. In making such a demand, the addresser implicitly recognizes the addressee's capacity to accept or reject the claim and expects the addressee to regard the demand as morally valid. [186] This reference to "a moral community" highlights a crucial element absent in both the interest and will theories, as well as in the kind-desire theory.[187]

To level the playing field, however, let's compare Darwall's theory with the will and interest approaches in the context of contractual obligations; this will permit a deeper exploration of how Darwall's framework addresses the relational and moral dimensions that were overlooked by its predecessors. As previously mentioned, Darwall views rights not as abstract entitlements but as demands grounded in interpersonal, rational relationships, where negotiation and mutual respect for claims play a critical role. However, Darwall does not view rights either as merely contingent entitlements: if I am justified in asking someone standing on my foot to move, I possess the moral standing to make that request—a legitimate right not to be stepped on—even if I choose not to exercise it. Darwall, thus, thinks that those who make agreements already possess prior demand-

[186] In this sense, the secondpersonal-dimension of rights not only establishes authority but also creates a specific normative relation that transcends general reasons for action. We will return on reasons and relational normativity in a constructivist framework in the next chapter.

[187] Darwall is explicit in stating that the moral community, as he understands it, is not any actual community of real human beings, but rather a "regulative ideal" that we use to make sense of our ethical thought and practice (Darwall, 2007, pp. 64–67). He also later put it as follow: "when we blame someone, we add our voice to, or second, […] a demand that we must presuppose is made of everyone by the moral community or a representative person as such." (Darwall, 2013c, p. 37). I will argue in the next section that Richardson conceptualization of the moral community is more systematic.

rights against one another, as an agreement would not be possible if they were not "already obligated to one another" (Darwall, 2006, p. 200).[188]

This position balances the need to protect individual autonomy with the recognition that in an intersubjective context, rights often require accommodation and compromise, making it a more practical approach in dealing with actual conflicts. Moreover, it avoids the problem of over-inclusiveness that can plague interest theory by being more selective about what constitutes a right. Not every interest gives rise to a right; only those interests that can be the basis of legitimate claims are recognized. This prevents the unchecked proliferation of rights and ensures that the concept of rights remains coherent and manageable.

Moreover, Darwall's theory addresses some of the challenges that arise for will theory, particularly in explaining the nature of power-rights and in accounting for the rights of individuals who may be unable to exercise their rights, such as minors or those with severe disabilities. Darwall suggests that even if certain demand-rights are not actively exercised, the agent still possesses them *in nuce*—that is, they retain the potential to make the claim, recognizing their right in principle. Additionally, while will-theory faces difficulties in explaining privilege-rights, where the right-holder has the discretion to choose whether or not to exercise their right without imposing corresponding duties on others, demand theory can maintain that some rights can be waived. For example, a teenager may waive their right to privacy by allowing their father to enter the room without knocking, if this does not disturb her. At the same time, demand theory maintains that some rights—such as the right to bodily integrity—are inalienable, meaning they cannot be waived or forfeited in the same way, maintaining a certain degree of flexibility.

To sum up, the main strength of Demand Theory lies in its ability to ground rights and obligations in the relational and moral practices of holding one another accountable; neither Will Theory nor Interest Theory explicitly addresses this interpersonal dimension, which is central to understanding how rights function in actual human practices. By incorporating this second-

[188] In the footnote of the same pages, Darwall makes the point clear by stating: "Just as God cannot obligate us by his commands without it already being the case that he has second-personal authority with respect to us, so also would it be impossible for two individuals to obligate themselves especially to one another by an agreement unless they already had the second- personal standing so to obligate one another in this way" (Darwall, 2006, p. 200).

personal dimension, demand theory bridges the gap between the abstract autonomy of Will Theory and the paternalist tendencies of Interest Theory.[189]

1.3.3 Richardson's Specificatory Theory

While agreeing with Darwall on the importance of mutual recognition, Henry Richardson places greater emphasis on the normative aspects of the moral community, which he defines as "the community of *all* individual person", beings who can wrong or be wronged (Richardson, 2018, pp. 7, 133). Richardson also more explicitly addresses the potential for individuals to reach a consensus about general moral norms, stressing the need for deliberative dialogue. To systematize this process, Richardson introduces a novel specification theory of dyadic rights and duties, grounded in the assumption that the moral community has the authority to establish new moral norms to address morally significant concerns.[190] This account, which Richardson terms precisely the "specificatory theory of dyadic moral rights and duties"(Richardson, 2018, p. 261),[191]

[189] Forst, like Darwall, rejects both will-based and interest-based theories of rights in his account of basic rights (Forst, 2016). He understands basic rights as "mutually justifiable claims" that protect individuals from legal, political, and social domination (Forst, 2016, p. 7). Crucially, these rights define individuals not only by securing their freedom from arbitrary interference, but by affirming their status as free and equal members of a normative order — that is, as co-authors of the legal and social structures that govern them (Forst, 2016, p. 8). As such, Forst's account already gestures toward a broader notion of accountability that extends beyond mutual recognition, aligning in key ways with Richardson's approach, as will be discussed in the following pages.

[190] Richardson focused on the specification of norms in one if his early works (1990), before focusing specifically on the specification of dyadic rights and duties. In this early paper, Richardson proposes specification as a "third way" between traditional models of application and balancing, as it involves qualitatively modifying norms to adapt them to concrete cases rather than merely weighing them quantitatively or intuitively. The innovation of this approach lies in its treatment of moral disagreement (rather than abandoning it, a conflicting norm is made more specific to resolve the tension) and in offering an alternative application method to overcome the theoretical deadlock between consequentialist and non-consequentialist approaches. "This hybrid model has served the very important and constructive function of convincing many that we need not cast aside a nonconsequentialist attachment to moral principles simply on account of counterexamples to each principle taken absolutely-for these can all too readily be generated from our considered judgments about hypothetical and real examples. Shifting from an "absolute" to a "prima facie" interpretation of moral principles allows the nonconsequentialist to survive in the face of such objections."(Richardson, 1990, p. 309).

[191] Richardson was among the first to introduce the terminology, and he deliberately avoided presenting his view as Specificationism, so as not to portray it as a fully formed "-ism" or comprehensive moral theory (such as utilitarianism or contractualism). Instead, he characterizes his approach as a method or process of refining moral norms to address indeterminacy: the adjective "specificatory" thus describes the function of this process—what specification does—rather than denoting a self-contained doctrine. Nonetheless, subsequent debate has drawn heavily on Richardson's insights while adopting a slightly different vocabulary. For instance, Elijah Millgram (2008, p. 732) speaks of Specificationism as a theory of practical reasoning opposed to Instrumentalism. Instrumentalism maintains that all practical reasoning is means–end reasoning aimed at achieving one's goals or satisfying one's preferences. It

posits the following: for individuals A and B, A has a moral duty to B to φ, and equivalently, B has a moral right against A that A φ, if and only if (i) B has a monadic moral right that, in the circumstances, entails that A has a duty to φ, and (ii) in virtue of (i), A possesses the moral power to specify A's duty to φ." "specification theory of dyadic moral rights and duties. It posits the following: For individuals A and B, A has a moral duty to B to φ and, equivalently, B has a moral right against A that A φ, just in case (i) B has a monadic moral right that implies in the circumstances that A has a moral duty to φ and (ii) on account of (i), A has a moral power to specify A's moral duty that A φ.[192]

In order to completely grasp the specificatory theory, we have to understand the role it plays in Richardson broader frameworks of Constructivist Ethical Pragmatism (CEP). In CEP which actions are right in a given circumstance is a function of the valid moral norms legitimately in place, where which ones are legitimately in place is partly a result of the moral community's tacit reflective efforts to work out a well-integrated and coherent understanding of the right and

thereby excludes the possibility of reasoning about ultimate ends: adopting final ends is treated as a non-rational process, neither correct nor mistaken. Specificationism, by contrast, holds that some practical reasoning involves refining abstract or vague ends into more concrete and specified forms, thereby making them suitable starting points for means–end reasoning (Millgram, 2008, pp. 732–733). Moreover, in more applied debates precisely on rights, the term Specificationism is generally used to contrast Generalism. If to say that a person has a moral right is also to say that others bear a corresponding duty, the central question becomes how the content and limits of such duties are to be determined. According to Generalism, moral rights derive from abstract principles that hold across all circumstances. According to Specificationism, by contrast, rights and duties must always be specified in light of particular contexts, and there are no principles valid *simpliciter* without exception. For example, on the generalist view, the principle that "every individual has a right to free speech" applies universally and establishes a corresponding duty on others not to interfere with that liberty. When conflicts arise—say, between free speech and public security—the generalist interprets them as requiring a balancing of principles, not as a matter of intrinsic specification. On the specificationist view, by contrast, the right to free speech does not mean "in all circumstances whatsoever"; rather, it must be specified so as not to include defamation, incitement to violence, and similar cases. For the specificationist, correlative duties are not rigid general rules to be balanced, but are already more finely determined by the surrounding moral and normative context. We will return to this issue in greater detail in §2.2.3, when addressing the problem of incommensurability in constructivist frameworks. For now, it is important to have clarified the use of the term specificatory in relation to these broader, though connected, debates. For more detailed discussion, see recent and more specialized contributions(Liberto, 2014; Moreso, 2012; Rettig, 2023)

[192] We can easily notice how Richardson definition directly introduces the theme of the nature of the rights; while his discussion engages extensively with legal scholars and significant jurisprudential debates, his focus remains on moral rights and duties, and he leaves "to another occasion, or to other hands" the question if the Specificatory Theory applies similarly outside morality, for example in law (Richardson, 2018, p. 262). Yet, as already seen in the introduction of this work, he is very categorical regarding the necessary distinction between laws and morality since the "moral community does not express its will in any formal way remotely resembling what's done in democratic settings" (Richardson, 2018, p. 179)

the good (Richardson, 2018, pp. 39, 264).[193] By noting that the authority of morality cannot arise in the same centralized way it does in institutionalized system like law, since the moral community is the open-ended set of all individual person, Richardson develop a three-stage conception of how the moral community can introduce new objective moral norms, that reduce morality's indeterminacy: 1) Authoritative specification 2) Convergent valuation and selection 3) Reflective ratification and acceptance.[194]

1) The first stage authoritatively specifies a moral norm in a way that reaches only particular people who are working out together how to best respond to a moral issue (the input). Here the relation between the agents is based on the moral responsibility embodied in the correlativity of dyadic right and duties, which arise from a moral division of labor wherein rights come to be specifically addressed (Richardson, 2018, p. 72). This dyadic specification express the ontological correlativity of rights and duties (Richardson, 2018, p. 111).[195]

2) The second stage moves from the particular moral authority two individuals have on each other to the kind of general moral authority that might establish a new moral norm. It does so by a selection among rival candidates innovations and a convergence of individual sufficient to establish a new social practices (Richardson, 2018, p. 135). Through trial and error, agents may develop a good sense of what counts as a reasonable specification, in a way that the new moral norm does not go against any of the extant norms, but instead filling in an indeterminacy that those norms leave open. (Richardson, 2018, pp. 143–144).[196]

---

[193] "Constructive ethical pragmatism is, roughly, the view that a) which actions are right in a given circumstance is a function of the valid moral norms legitimately in place, where which ones are legitimately in place is partly a result of the moral community's tacit reflective efforts to work out a well-integrated and coherent understanding of the right and the good; and b) where the valid norms legitimate generate a practical indeterminacy, which actions are rights may further be determined by additional efforts in that direction" (Richardson, 2018, p. 39)

[194] Richardson primarily refers to these stages as the input stage, convergent stage, and ratification stage (Richardson, 2018, pp. 21, 263). Additionally, he labels the first stage as the "felicity condition" and the latter two as the "legitimacy condition". (Richardson, 2018, p. 26)

[195] "As a specific addressee of the relevant (initially monadic) right, the duty-holder has been placed in the position of having to carry out intelligently a moral responsibility vis-à-vis some particular person. The fact that the duty-holder is called upon intelligently and conscientiously to promote or respect the other's right explains both why the duty-holder has a duty to the right-holder and why the right-holder has a right against the duty-holder, establishing the ontological correlativity of this duty and right. In this way, dyadic rights and duties arise."(Richardson, 2018, p. 111).

[196] Even if duty-holders and right-holders will have "some built-in incentive to disagree on the detailed specifications of the relevant moral material" each will also have a "good basis, in their self-knowledge of their

3) Once a moral norm passes the second stage, it indicates that its tacit acceptance by individuals is already widespread. What remains are two final steps: first, developing "a reflective awareness of the social convergence"—meaning that individuals recognize that others accept the norm; and second, achieving a collective reflective acceptance, where individuals become "mutually aware of each other's acceptance" of the norm, allowing us to say that we all know we accept the norm (Richardson, 2018, p. 177). Reflective acceptance does not require unanimity (since the moral community is generally not subject to voting or other procedural determination), but there are others four conditions.[197] Once a moral norm is ratified, it carries the implication of accepting it as a universally binding one.

This is the overall proposal of Richardson, through which we can reach a legitimate moral norm.[198] Let's focus for the present on the first two stages to see in detail what is the novelty of the Specificatory Theory that Richardson introduces regarding the relationship of rights and duties. The Specificatory Theory is designed to address challenges like those posed by Michael Thompson (2004) regarding how all persons can be united into one normative community which aspires to be open-ended (Richardson, 2018, pp. 28, 75–79).[199] Richardson's answer is that the moral community, structured by a web of correlative rights and duties, can encompass new members as long as they share the capacity for moral interaction (wronging or be-wronged) and thus being capable "of dyadically joining together hitherto-isolated groups of intelligent beings" (Richardson, 2018, p. 262). The fundamental axiom of moral equality is enacted by the duty-holder, having been

concerns and in their experience, for setting aside some specifications as simply silly or as nonstarters" (Richardson, 2018, pp. 143–144).

[197] "(i) Broad and inclusive, and broadly and inclusively recognized, reflective acceptance of the norm in question as a moral norm, (ii) which acceptance has been informed by broadly and inclusively available critical reflection and (iii) has arisen at least in significant part because of a similarly broad and inclusively acceptance that the norm was a reasonable response to a moral problem and (iv) has been informed by successful and reasonably sound efforts at reconstructing the process of the norm's becoming a global social norm as a process of reasoning about how to fulfill morality's requirements or how to carry out morally urgent tasks." (Richardson, 2018, pp. 263–264).

[198] On the relation between norms and rights we will return in the next chapter.

[199] "How we can conceive of a moral community of persons in which it is the case both that they are knit to one another via a web of directed duties and the correlative rights, and that the resulting community is open ended?"(Richardson, 2018, p. 28)

given a special moral responsibility as an equal moral person (Richardson, 2018, p. 261),[200] which permit the unity and inclusiveness of the moral community (Richardson, 2018, pp. 118–134).[201]

So conceptualized, the specification theory of rights is able, according to Richardson, to address the challenges faced by both interest and will theories. The problem with interest theory is that it reduces rights to the protection of specific interests—those deemed essential for an individual's well-being or a good life—thus reducing dyadic rights to monadic rights.[202] This reduction fails to capture the interpersonal and normative dimension of rights, which involves recognizing individuals as moral agents with agency, not simply as beings whose interests need protection, namely moral patient; while interest theory seems to assume this dimension in some cases, it does not adequately address it (Richardson, 2018, p. 91).[203] This critique also applies to Wenar's Kind-Desire Theory, which similarly fails to clarify how an individual's interests can ground a dyadic duty, leaving the connection between rights and duties in this framework unclear (Richardson, 2018, p. 92).

---

[200] "The moral power to specify one's duties that lie at the core of the Specificatory Theory are power that derives from the duty-holder having been given a special moral responsibility as an equal moral person. The fundamental axiom of moral equality underwrites seeing all persons has having the relationship to morality postulated by the participant conception of rules, in which each person, equally, is called upon, insofar as they are able, to approach their moral responsibilities intelligently, and so with an openness to reconsidering their formulation or point." (Richardson, 2018, p. 261).

[201] In order to fully achieve this step, we need to have both agents as morally competent in "working out together some of the specifics" about how the duty-holder should fulfill his duty. "If the duty holder is not competent to specify the duty, it is hard to imagine that the duty-holder is competent to fulfill it." (Richardson, 2018, p. 87).

[202] "The idea of wronging someone is a dyadic notion: one person wrongs another. It is to be contrasted with the monadic idea of acting wrongly." (Richardson, 2018, p. 66) While some authors are suspicious of the idea of a 'monadic' right and will be happy to simply discard it, believing that something only merits the title of a right if it already entails some duty since a right correlating with no duty is indistinguishable from something merely being valuable (Cruft, 2019, pp. 6–7), Richardson does not argue that monadic rights do not exist. He simply maintains that when we refer to claim-rights one of its main functions it's the specification of other duties, but there are also some types of addressing that are not dyadic. As Richardson puts it: "Although dyadic rights can be addressed, an addressed right need not be dyadic. Yet the explanatory connection built into this idea of a right's being addressed provides a crucial element of the social background of the Specificatory Theory, which presupposes such an explanatory link between its two main elements." (Richardson, 2018, p. 95).

[203] "The interest theory presumes dyadic form but makes no attempt to explain it" (Richardson, 2018, p. 91). This happen also because, according to Richardson, there is a (self-)confusion when interest theory makes reference to other people from the right-holder perspective, "mimicking" dyadic structure. "Instead of involving, say, duties to others, all they involve is what Kant called duties 'regarding' others." (Richardson, 2018, p. 76). This is a point connected with the distinction between being "patient" or being a "counterpart" of a duty. On the distinction between agent and patient, and how the concept of patient produce confusion, see also O'Neill (1996, pp. 94–95).

The will theory is by contrast non-reductive: it explains dyadic rights and duties on the basis of the high order dyadic pair of powers and liabilities, putting the focus on the right-holder. (Richardson, 2018, pp. 92–93).[204] A promise is owed to the promisee by reference to the fact that the promisee has the power to waive her claim to performance. However, if I observe that Lucas is exploiting Daniel's kindness by repeatedly breaking his promises, even though Daniel is accommodating and always forgives him, I can still recognize a form of injustice in their relationship. The will theory fails to address this injustice because it cannot account for the concept of inalienable rights. According to will theory, the defining feature of a rights-holder is their moral power to control their rights. However, those who possess inalienable rights lack this core power of control, leaving the theory unable to explain the moral significance of such rights, as Wenar (2013, p. 226) aptly underlined.

The Specificatory Theory fix these problems by employing a similar nonreductivist strategy,[205] but "on the side of the duty-holder" (Richardson, 2018, p. 92) and is thus less prone to troublesome counterexamples and has more metaphysical parsimony (Richardson, 2018, p. 261). Moreover, the Specificatory Theory surpasses the notable shortcomings of Will Theory in accommodating individuals who cannot fully exercise autonomy by grounding obligations in the reasoned norms of a community, which can recognize and advocate for the rights of those who lack the capacity to articulate demands themselves. While Richardson admits that it is preferable for two competent individuals to resolve moral issues symmetrically when feasible for at least three types of reasons (epistemic, recognition, and pragmatic),[206] unavoidable asymmetries often

[204] "This explanatory potential arises from an important conceptual asymmetry between dyadic rights and moral powers. Whereas there are monadic rights and duties, which are not held against or owed to anyone, moral power are necessarily dyadic in a way that supports Hohfeldian correlativity: normative powers to affect someone's Hohfeldian incidents are equivalent to liabilities in that person. 'A moral power that is not a moral power over someone' is oxymoronic" (Richardson, 2018, p. 93)

[205] "Rather than defining the right as a function of goodness, as consequentialist theories do, or understanding the constrains of right as being fixed independent of goodness, as deontological theories do, CEP is a nonreductively integrative moral theory" (Richardson, 2018, p. 264). We will revisit the non-reductive element of constructivism in the next chapter, specifically in Section 2.3.

[206] "One reason they ought to do so is epistemic: the right-holder is likely to have a perspective on the relevant moral considerations that well complements the addressee's. Conscientious fulfillment of the responsibility in question thus calls for hearing the right-holder's perspective. Another reason is that doing so is called for by due recognition of the other as a moral equal whose morally relevant interests are at stake. Respect for the other as a moral equal calls for hearing the other's point of view, if the other can articulate it or if it can be expressed by a proxy or surrogate. A final reason is pragmatic. Working these things out together will help coordinate these two agents, getting them on the same page morally" (Richardson, 2018, p. 112).

arise in cases involving right-holders who are not morally competent and lack a proxy to represent them, and an adequate moral theory needs to be able to address these cases (Richardson, 2018, p. 112).

To reconnect with our broader aim, we can notice that inside the constructivist framework Richardson offers a unique systematic analysis of the functions of rights. While both Darwall and Richardson emphasize the interpersonal dimensions of rights, their approaches diverge in focus: Darwall centers on the direct accountability inherent in the second-person standpoint, highlighting the moral dynamics of addressing claims and responsibilities between individuals, while Richardson extends the scope to a broader communal reasoning process, emphasizing how collective deliberation within a moral community can specify and uphold rights and duties. Notably, while Richardson himself underscores "the distance between dyadic moral relations and second-personal address" (Richardson, 2018, p. 83n),[207] it remains entirely feasible to view the Specificatory Theory of rights as compatible with Darwall's framework, provided that a demand can be understood as the specification of a right in correlation with a duty. Both theories, despite their distinct emphases, can contribute to a richer understanding of the normative and interpersonal functions of rights and duties in a constructivist sense.

Regarding the priority of duties, Richardson partially agrees with O'Neill that talk of rights without specifying counter-duties is what gives "human rights talk a bad name" (Richardson, 2018, p. 94) but also tries to resist O'Neill suggestion that if the specification of right is not a priori or noncontingent, then it is merely conventional and lacks universality and objectivity (Richardson, 2018, p. 103n). Richardson's approach suggests that rights are not merely passive reflections of pre-existing duties but are also the simultaneous outcomes of active deliberation about what individuals are owed. This process of moral deliberation establishes both the rights individuals possess and the duties others owe concerning those rights. In this bidirectional relationship, duties and rights are co-dependent, working together to shape a just and moral society. This process is

[207] In another note Richardson also writes the following: "It has recently become popular to attempt to explain the usefulness of the idea of a directed duty by reference to the idea that the right-holder gains a special "standing" to complain of nonperformance. See, e.g Darwall (2012). Since the relevant idea of standing is itself quite elusive, perhaps an adequate rebuttal to this suggestion is to ask why such standing is not already adequately explained, and without invoking true directedness ('you had a duty to me to φ!') by reference to being a duty's patient ('you had a duty to φ me!')." (Richardson, 2018, p. 75n).

characterized by both a universal (morality objective core) and a contingent aspect (the pragmatic ratification by the moral community), and it ensures that human rights remain rooted in the shared capacity of all persons to participate in moral deliberation, transcending individual differences and particularities. In this way, moral rights are neither static nor arbitrary but centered on both universal and actionable duties shaped through the specification of the correlative rights (Richardson, 2018, p. 95).[208]

Therefore, in sum, by understanding rights and duties as mutually constitutive and actively specified, Richardson's approach can be integrated with the broader discussions of Darwall and O'Neill. This integration paves the way for a unified constructivist conception that preserves both the objective core of moral rights and the practical applicability of basic rights we seek.

1.4 A PRELIMINARY CONSTRUCTIVIST DEFINITION OF RIGHT

Let us briefly recapitulate the main points established in this first chapter. Firstly, the question concerning the fundamental nature of rights—whether they are primarily moral or legal—is, in fact, a misleading or ill-posed question. It makes little sense to pose the issue in these binary terms, since morality and legality represent two distinct yet interconnected domains of normativity. Constructivist thinkers, despite their varying emphases, agree that there exists an important link between law and morality, and this is why they reject the separation thesis, which treats legal and moral normativity as entirely independent, in favor of a model that emphasizes their normative congruence and overlap (Forst, 2002, p. 67).[209]

[208] "Monadic human rights may be said to 'exist' in an abstract way if they ought to be promoted and respected. Compatible with thinking this, however, we may […] hold that rights that are not addressed lack 'real' existence. Until it is specified who is responsible for protecting and respecting them, their existence remains of merely theoretical, rather than practical importance." (Richardson, 2018, p. 95).

[209] However, as already mentioned, a constructivist is committed to rejecting the separation thesis—the idea that law is not based on any moral norms—while still upholding the ethical neutrality thesis, which asserts that law should not be based on any specific ethical conception of the good. As Rainer Forst explains: "Liberal principles of law constitute not a particular form of life, but a system of justified norms that, to be sure, is not without ethical demands (and not entirely without ethical content) but is not justified on ethical grounds. Only reciprocally and generally justified law can be ethically pluralistic, open, inclusive, and legitimately binding" (Forst, 2002, p. 67). Law is neutral insofar as the threshold of reciprocity and generality are not themselves based on ethical values, but are moral universal norms (Forst, 2002, p. 234).

This explains why the concept of unjust laws is both meaningful and important for constructivist theory. Why, instead, is it less intelligible to think of an unjust right? The main logical form of right, that of a claim, does not directly exclude this possibility, since it is perfectly reasonable to imagine someone claiming a right to discriminate against minorities; this is due to the fact that, per se, the logical form of a claim only entails the existence of a duty with a certain structure, but not necessarily a normative structure. Constructivists accept the logical correlation of (claim)-rights with (directed) duties, but they do not assert that all duties have correlative rights (as in cases of imperfect duties). This does not mean that the concept of duty takes precedence over that of rights (O'Neill, 1996, p. 135) but rather that the two are not always in a strictly reciprocal relation. When rights and duties do correlate, however, they form a bidirectional and interdependent structure, rather than a hierarchical or causal one: neither the right nor the duty precedes the other, but they are co-defined through their relational function.

This function is not adequately captured by classical theories: not merely a tool of choice and autonomy as in Will Theory, nor simply a protector of individual desires and well-being as in Interest Theory. Rather, rights are best understood within a constructivist account as demand of mutual accountability (Darwall, 2006, p. 142). To say one has a right, is to say that one is in a normative position to hold others accountable to justified or legitimate claims within a shared social, legal or moral order. Furthermore, this social function can be expanded from accountability based on mutual recognition to accountability based on the legitimate conditions defined by a broader community that specify relations of rights and duties about what individuals are owed. In this bidirectional relationship, duties and rights neither static nor arbitrary but co-dependently specified (Richardson, 2018, p. 95). This specification is what enables us to distinguish between a just and an unjust right—or, more precisely, between a legitimate and an illegitimate specified demand. For example, a slave-owner might wish to claim the right to possess a slave, but such a claim cannot constitute a legitimate demand in this qualified sense.[210]

[210] It might be argued that if a law permits slave ownership, then such a demand is, by definition, legitimate. This raises the deeper question of the distinct normativity of law and morality, and the different forms of legitimacy each entail—a topic I will briefly touch on in the following paragraphs and examine in greater detail in the next chapter. However, in this example, it seems evident that in any genuine rule of law system slave ownership would be impossible, as slavery fundamentally violates one of the core principles the rule of law is intended to uphold, legal equality, which requires that all individuals be treated equally before the law. Historical examples of legal systems that claimed to operate under the rule of law while allowing slavery—such as the United States prior to the Civil

Building on these key points, we can sketch a concise, constructivist-oriented definition of rights. Without, for the moment, distinguishing between moral and legal rights or the forms of legitimacy they imply, a right can be generally understood as: a social claim (which can be abbreviated as demand) that specify and is specified by corresponding duties on others. Most constructivists would accept this definition, even if they do not acknowledge or consider outdated some of the preceding distinctions and classifications.[211] The definition can be formalized as follows: "To assert that A has a (legal or moral) right against B is to affirm that A possesses a (legal or moral) demand on B, and thus, there exists a correlative duty for others to refrain from violating this claim." For instance, a teenager's right to privacy specifies a social claim against someone entering their room without knocking, and others have a corresponding duty to respect this claim. This duty does not necessarily arise directly from her specific claim but may instead be specified by the duty-holder, who reflects on the situation and determines and specifies their obligations accordingly.

Some could argue that this definition is too minimalist, and that it is crucial to provide additional information and more characteristics in order to have a truly useful conception of right. For instance, consider the following comprehensive definition:" To say that A has a (legal or moral) right against B is to say that A has a (legal or moral) claim on an act or forbearance from B-meaning that, should B fail to so act or to so forbear, it would be (legally or morally) justifiable for A to use coercion to extract either the act or forbearance from B, or compensation in lieu of it" (Becker, 1980, p. 1198). Now, if we substitute the notion of claim with that of demand, to express not only the logical form but also the specific social function that rights endorse, this characterization seems to suggest all the elements necessary for analyzing the concept of a right as it is currently used and would still likely be endorsed by most constructivists. Specifically, it suggests that a complete account of a particular right involves:

War—only demonstrate that such systems upheld the rule of law in name only, having excluded entire classes of people from legal protection.

[211] Especially regarding the Will-Interest debate, which is considered to be outdated by many of them, as we noted in the footnote n.160 on Ripstein. Constructivists are interested to underline that is possible to affirm some legitimate claim-rights in a non-merely arbitrary way: whether legitimacy derives from objective interests or from the ability to determine what others may and may not do (will) is not necessarily a question they are interested in answer, but I have tried to show an alternative that has been proposed by some of them (Forst, 2016; Richardson, 2018).

1. Specifying the right holder(s)—that is, identifying the individuals, beings, or entities that possess the claim.
2. Identifying the right regarder(s)—that is, determining the "others" against whom the right holder has a claim.
3. Specifying the acts or forbearances required—clarifying the actions or abstentions expected from the right regarder(s); what they may or may not do; or, in the case of "regarderless" rights, the benefits to which the right holders have a claim.
4. Defining the nature of the rights-relationship—determining the type of claim the holder has on the regarder, or the type of claim the holder has on some benefit.
5. Establishing the conditions for a violation of the right—identifying the circumstances, if any, that constitute a right regarder's culpable failure to act or forbear as required.
6. Determining the conditions under which a violation of the right (though culpable) is excusable.
7. Specifying the compensation, the remedies for both excused and unexcused violations of the right.
8. Identifying coercive measures for enforcing the remedies
9. Determining the agent(s) authorized to extract the remedies.
10. Justifying the right as defined above—presenting the reasons that warrant the conclusion that the right holder(s) do, in fact, have such a (legal or moral) claim.

However, upon a closer look, these ten features demonstrate the potential pitfalls of describing every right as claims on an act or forbearance from another. For instance, we cannot universally assert that “compensation” is always interpreted literally, since in some cases a simple apology may suffice (Becker, 1980, p. 1199). Additionally, “coercion” may not always entail literal force (in certain situations, a verbal demand may be appropriate), and the right holder may not always be the one authorized to apply coercion (for instance, only a law enforcement officer may do so). The minimalist definition that we provided, by avoiding these qualifications, remains broad enough to be accepted by the majority of constructivists.

This is just the first building block towards our definition: from this austere understanding of rights in general, we can further specify the types of rights we are interested in, not only in terms of their nature, forms, and functions but also according to the normative context in which they are situated and the specific aspects of reality they address. This is our goal in defining human rights in a constructivist way: in Chapter Two, we will explore the normative aspects of human rights by comparing them with other normative concepts, while in Chapter Three, we will clarify the specific beneficiary that human rights aim to protect (the human being itself).

Yet, we can already briefly notice what seemed a recurring theme from the analysis, particularly in the final sections of this chapter addressing Darwall and Richardson: the constructivist perspective increasingly appears unable to avoid referencing morality in order to reach a complete understanding of the notion of rights and especially of human rights. In order to provide a thorough examination of the distinct bindingness of moral and legal rights, merely asserting that they are not separable but rather overlap, does not suffice. Instead, it is essential to elucidate that, from a constructivist perspective, the normative context of moral rights is, in some qualified sense, broader and axiologically antecedent (Forst, 2024).[212] The normativity associated with moral rights will be the beginning of our analysis in the next chapter.

[212] A considerable number of non-constructivist scholars concur with this thesis, particularly, but not only, those who align more closely with the overarching proposal previously referenced, such as pragmatists, contractualists, and demand theorists. Gilbert for example express the relation as follow: "A key distinction between moral and legal rules has to do with their normative implications. Moral rules, as such, have normative implications. Legal rules do not. Moral rules are thus in an important way more powerful than legal rules" (Gilbert, 2018, p. 46). While many constructivists, like Forst, while says that the general notion of normativity is broad and include the normative implications of legal rules expressed in sanctions and practical necessities, they typically maintain that the normativity of law is derivative and, in some important sense, presupposes the normativity of morality. Nonetheless, Gilbert is highlighting this distinction between moral and legal rights to assert that demand rights is a distinct third kind of rights, something that the majority of constructivists are not likely to endorse.

CHAPTER 2
FURTHER NORMATIVE CONCEPTS

Ask yourself, what is a reason? It is not just a consideration on which you in fact act, but one which you are supposed to act; it is not just a motive, but rather a normative claim, exerting authority over other people and yourself at other times. To say that you have a reason is to say something relational, something which implies the existence of another, at least another self. It announces that you have a claim on that other, or acknowledges her claim on you. For normative claims are not the claims of metaphysical world of values upon us: they are claims we make on ourselves and each other.

(Korsgaard, 1993, p. 51)

By examining rights through the constructivist lens, we started to move from a provisional definition (rights are claims) to a more articulated one: rights are legitimate (or normative) demands that specify and are specified by corresponding duties on others. However, what makes a claim legitimate or normative? What is and where does the normative force of these claims come from? Strictly speaking, one cannot have a claim-right to perform an action, as asserting claim-rights does not yet imply assertions of permissibility or normativity: it only entails the existence of a duty with a certain structure, and not a duty with a certain force (Sreenivasan, 2010).[213] For instance, returning to the example from paragraph 1.2, Daniel may have a claim-right against interference with him giving his car to Lukas, even if he has an obligation to give it to someone

[213] Recall that the logical form of a claim right is the following: *A* has a claim right against *B if and only if* *B* has a duty toward *A*. In this definition it is the "direction" of *B*'s duty (that it is owed to *A*) which correlates *B*'s duty to *A*'s right; nothing is said in the definition about the duty's strength. Why "directed duties" have greater normative priority remains an open question: the only broad consensus is that they do (Wenar, 2023).

else (perhaps he promised it to his son) or if it would be morally wrong to give it to Lukas (perhaps Lukas intends to use it for a robbery). Alternatively, consider a different example: I may have a right not to give the money I spend on luxuries to the needy; in this sense, I have a claim-right against others forcing me to do so, but this claim-right, per se, does not supply a justification for my decision to keep the money. Here the rights just affirm (or secure, through legislation) a sphere of non-interference, but they do not imply that my choice within that sphere is morally good, wise, or even defensible.[214]

While for many scholars, an essential characteristic of rights is precisely their capacity to permit actions that one ought not to do (Dworkin, 1977, p. 188; Raz, 1979, p. 274),[215] from a constructivist perspective, where rights are primarily understood to function as legitimate specificatory demands, the notion of unjust right —or more broadly, the right to act wrongly— seems a self-defeating notion which undermines the very normative function that rights are meant to serve. To address this apparent contradiction, we need to first clarify how the normativity of moral rights and legal rights diverge. This requires a broader understanding of how constructivist theories conceive of normativity itself (section 2.1), and how a concept like that of rights relates to other normative notions such as reasons and values (respectively sections 2.2 and 2.3). By mapping these conceptual connections, we will begin to see that the apparent divide between moral rights and legal rights is not a matter of their fundamental nature, but rather but rather in the way they are applied and operationalized within distinct normative frameworks. This clarification will allow us to specify with greater precision what the notion of a right entails within a constructivist framework (section 2.4).

[214] I have used these examples to illustrate once again how Hohfeldian incidents interconnect, and how constructivists tend to conceive of the *claim-right* as the overarching form under which the others can be analyzed. Depending on how the example is formulated, the right in question may appear as a liberty, a claim, or even an immunity. For instance, if I say that I have no duty to donate my money to the poor, I am pointing to a liberty-right; if I say that others have a duty not to coerce me into donating, I am invoking a claim-right; and if I add that the state itself lacks the authority to impose such a duty, then I am appealing to an immunity-right.

[215] This general feature of claim-rights is recognized and accepted from both jusnaturalists and juspostivists. Dworkin, for example, explains that: "when I say you have the right to spend your money gambling, though you ought to spend it in a more worthwhile way […] it would be wrong for anyone to interfere with you even though you propose to spend your money in a way that I think is wrong" (Dworkin, 1977, p. 188). Similarly, Raz claims that "[t]o show that one has a right to perform the act is to show that even if it is wrong he is entitled to perform it" (Raz, 1979, p. 274).

2.1. NORMATIVITY AND AGENCY

Broadly speaking, normativity characterizes principles, rules, standards, and disciplines that guide actions and beliefs with some degree of authoritative force, making them more than mere suggestions, but *norms* (from which the term originates).[216] Norms stand in contrast to descriptions: while both are representations of actions and beliefs, categorizing something under a description defines its type, whereas categorizing it under a norm establishes its legitimacy, what we should do or believe, the guiding "ought" that shapes judgments.[217] However, more specific and differentiated definitions abound across various fields, not only in moral and political thought but in other areas as well.[218] Even within metaethics, different subgroups propose various relationships involving normativity: some emphasize a constitutive link between normativity and motivation, others between normativity and intentions to act, or between normativity and agency

[216] Here, I am not yet presenting a specific constructivist conception of normativity, which can also vary significantly among different constructivists. Instead, I am offering a general definition that could be accepted by a broader range of scholars, even though they may characterize the concept of norms in different ways. To illustrate how constructivists conceptualize the concept of norms, here is Forst's specific definition: "A Kantian constructivist view entails two different kinds of normative arguments or two kinds of normativity: first, the normativity of the principles and ideas of practical reason, to use Rawls's language – that is, on my account, the principle of reciprocal and general justification and the moral notion of free and equal persons as equal normative authorities with a right to justification; and, second, the normativity of the norms (or 'laws') generated by the constructivist or discursive procedure, be it the categorical imperative or a notion of free and equal discourse." (Forst, 2016, p. 13). Richardson has a weaker definition of norms which he thinks of "simply as a principle – an item with propositional content – which has at least potential normative significance" (Richardson, 2018, p. 50)

[217] This is why discussion of normativity are usually tackled by distinguishing *theoretical reason* (concerned with reasons to believe) and *practical reason* (concerned with reasons to act). We will return to this point in section 2.2.3, but here it is worth briefly restate again that many constructivist scholars prefer to frame this division in alternative terms; rather than emphasizing a sharp separation, they often highlight the continuities and shared structures between theoretical and practical reasoning—suggesting that both are guided by similar principles of justification and responsiveness to reasons (Korsgaard, 1986; Forst, 2021; Bagnoli, 2022c). For example, according to Rainer Forst's distinction between contexts of justice, practical reason is a justificatory reason and refers to the capacity to provide reasons appropriate to specific normative contexts. This entails aligning one's actions and judgments with moral principles (e.g., "One shall not lie"), ethical rules (e.g., "Greet people when you meet them"), legal standards, (e.g., "Drivers must stop at red lights"), and political practices (e.g., "Citizens should exercise their right to vote"). "The person at the center of a theory of justice is not just an ethical person, a legal person, citizen, or moral person, he or she is all of these simultaneously in a different way: he or she is ethically, legally, politically, morally autonomous" (Forst, 2002, p. 240).

[218] For example, see Glüer, Wikforss, and Ganapini (2023) for how the concept of normativity is central in the philosophy of language, or refer to Steinberger (2022) for the normativity of logic.

(Copp & Morton, 2022). Constructivism falls into this latter group: we are agents, and as agents we are continuously engaged in performing actions. The key question then becomes: in what sense are our judgments about actions normative for us?

In *The Source of normativity*, Christine Korsgaard says that the moral philosophy of the modern period can be read precisely as a search for the source of normativity and she distinguishes between four different conception on normativity: 1) Voluntarism 2) Realism 3) Reflective Endorsement 4) Kantian constructivism (Korsgaard, 1996b, pp. 18–21). For voluntarists, obligation derives from the command of someone who has legitimate authority over the moral agent and so can make laws for her (you must do the right thing because God or a political authority commands it); for realists, obligation derive from true moral claims (a moral claim is true if there are intrinsically normative entities or facts which it correctly describe); for philosophers that subscribe to reflective endorsement, like Hume, Mill and Williams, obligation is grounded in human nature (the central point is not that normative claims are true, but that they are good for us); finally, for Kantians constructivist the source of the normativity lies within the agent's own will: moral norms are the laws of the agent's will, representing claims the agent is prepared to impose on herself in an act of self-constitution (Korsgaard, 1996b, p. 94, 2009, p. 27).

In Korsgaard's view, Kantian constructivism most effectively explains the source of normativity by illustrating how it arises from a constitutive relationship between agency, rationality, and morality: actions are not mere responses to external forces, but are instead directed by reasons and principles we endorse as our own. Through our choices, we continually shape and affirm our identity as an autonomous agent, but at the same time our acting is constrained by certain principles fundamental to any rational being. [219] These principles are both *descriptive* (they *describe* what rational agents do) and *normative* (they *prescribe* what rational agents ought to do).

[219] The concept of being both free and constrained is central to self-constitution, but it does not represent a strict logical consistency, just a practical one. This is why Korsgaard emphasizes "practical identity"—the set of roles, commitments, and values that define who we are (Korsgaard, 1996b, p. 101). Normativity, then, arises from this practical identity, as being true to ourselves (e.g., as friends or citizens) requires us to act in ways consistent with these roles. However, Korsgaard later argues that in order to remain practically coherent with the particular identities we choose, we must first and foremost endorse a more fundamental identity: that of a human being as a "reflective animal". "Rational reflection may bring you to discard a way of thinking of your practical identity as silly or jejune. What is not contingent is that you must be governed by some conception of your practical identity. For unless you are committed to some conception of your practical identity, you will lose your grip on yourself as having any reason to do one thing rather than another - and with it, your grip on yourself as having any reason to live and act at all. But this reason for conforming to your particular practical identities is not a reason that springs from one of those particular

While descriptions cannot be violated—since they are simply true or false—norms can be disregarded or refused.[220] But, as Korsgaard argues, rejecting the normative element is not a neutral stance: it destabilizes our very status as rational agents. In turning away from normativity, we fracture our practical identity, weakening our capacity to act purposefully and with integrity (Korsgaard, 1996b, p. 235).[221] Self-constitution then becomes not the affirmation of autonomy, but a form of self-alienation (Korsgaard, 1996b, p. 103, 2019, p. 114). This is why skepticism about moral obligation is incoherent: skeptics cannot deny the authority of morality while still adopting a deliberative stance (Korsgaard, 1986, 2009, pp. 32–33).[222]

While the debate on normativity initiated by Korsgaard places a strong emphasis on the first-person deliberative stance,[223] framing it as the perspective through which an agent's moral

---

practical identities. It is a reason that springs from your humanity itself, from your identity simply as a human being, a reflective animal who needs reasons to act and to live. And so it is a reason you have only if you treat your humanity as a practical, normative, form of identity, that is, if you value yourself as a human being." (Korsgaard, 1996b, pp. 120–121). The importance of valuing humanity within a constructivist framework will be addressed in paragraph 2.3.2.

[220] For example, if a scientific law such as gravity does not correctly account for an observed phenomenon, the description is not "violated" but rather shown to be false, incomplete or in need of revision. By contrast, one can genuinely violate a norm by breaking a promise.

[221] "I am not saying that it is obvious that these rules are constitutive of these activities and that no argument needs to be made. I am making a point about what such arguments, when they are made, aim to achieve. If the argument of the previous section works, then the requirement of universality is in this way constitutive of willing" (Korsgaard, 1996b, p. 235). We will return to the concept of universalizability, which is considered by the majority of Kantian Constructivists the most important constitutive norms of reason, in paragraph 2.2.4.

[222] This inescapability is "deliberative" because it essential to the structure of agency itself. If moral obligations limit or constrain personal goals, it is because these goals would ultimately undermine the agent's own coherent self. While the norms of rationality are constitutive, they also remain normative: though they can be violated, doing so risks losing any clear, agent-centered perspective on action. We will return in detail on the debate regarding the inescapability of agency for constitutivism in chapter 5, paragraph 5.3.1. It is important, however, to briefly anticipate that the contemporary debate on inescapability is no longer centered on the notion of moral agency; with the exception of a few Korsgaardian defenders (Schapiro, 2021), this view has largely been abandoned (even Korsgaard herself has moved away from presenting the issue strictly in these terms, though her position remains somewhat ambiguous). The focus today is instead on the broader notion of practical agency, understood as the capacity of reflective beings to act for reasons. David Velleman is often credited with having shifted the discussion to this more precise and less morally loaded framework (Velleman, 2000a, 2009). See the subsequent footnote n.227 for how Velleman distinguishes two senses of inescapable.

[223] This may seem counterintuitive, since even in the opening citations of this chapter Korsgaard explicitly speaks about the relational nature of reason. And yet, if we reread carefully the passage we notice her stating that "a reason is to say something relational, something which implies the existence of another, at least another self." (Korsgaard, 1993, p. 51). Reasons are relational because "it takes two to make a reason" (Korsgaard, 1996b, p. 137), but since the two are for Korsgaard two elements of reflective conscience, "the thinking self" and the "acting self", her conception of relationality remains ultimately monological (Bagnoli, 2007, p. 116).

obligations are constituted, not all constructivists agree that the full source of normativity can be captured within this first-personal framework. Some underline the importance of the second-person stance, which emphasizes the relational aspect of obligations, which arise from normative claims people make on each other, establishing mutual constraints in deliberation through their shared normative relations (Darwall, 2006).

Further expanding this trajectory, Carla Bagnoli suggests that there is an interesting sense in which deliberation is about reaching the first-person plural stance (Bagnoli, 2007). This collective stance involves reasoning not merely as individuals but as part of a group agent, where moral reasoning transcends individual perspectives to reflect a more collective, or group-based, source of normativity. Bagnoli critiques Korsgaard's conflation of the reflective endorsement of moral obligations and the self-constitutive nature of normativity (Bagnoli, 2021c, p. 406);[224] moral norms are authoritative not solely because individuals endorse them but because they are constitutive to rational agency itself and this applies to every ration agent. The normative source of moral obligations is thus not (only) the moral authority in the first-person singular but that there are constitutive norms that bind all in the first-person plural (Bagnoli, 2021c, p. 407).[225] Normative obligations thus emerge as a dialogical reasoning process rather than merely an individual reflective endorsement.[226]

---

[224] "Korsgaard seems to assume that the argument based on reflective endorsement and the argument from self-constitution are ultimately equivalent, at least functionally, in that they both aim to establish that moral obligations are inescapable. However, the argument from reflective endorsement does not establish by itself that moral obligations are constitutive of the self; it only establishes that they are inescapable in first-person deliberation, at least on two conditions (i) if reflection is governed by the norm of universality, and (ii) if the norm of universality affords moral obligations. By itself the individual capacity to fully endorse or reject a proposal of action does not assure that moral obligations should have a unifying or grounding role in the constitution of the self" (Bagnoli, 2021c, p. 406)

[225] "Differently than the issue of the normative source, the issue of subjective authority does arise in the stance of deliberation. But it is questionable that deliberating agents are bound by moral obligations only insofar as they are capable of reflective endorsement. That is, the process of endorsement is successful in warranting normative authority only because and insofar as the reflective stance is governed by norms respectful and protective of humanity. However, if the self-constitutivist argument does its job, then the mechanism of reflective endorsement is an idle addition. For the self-constitutivist view establishes that there are constitutive norms of rationality that justify moral obligations as grounded on agential features, and by this means it purports to explain their normative and subjective authority." (Bagnoli, 2021c, p. 407).

[226] "The self-constitutivist view aims to establish a stronger connection, and say that wrongdoing is self-defeating. If the self-constitutivist argument is successful, then, any agent has a normative reason not to commit wrongdoing. Yet, this is not to say that the reason why wrongs are wrongs depend on the way they affect one's agency and integrity. […] To conflate these two issues is to conflate the normative source of moral obligations with the subjective authority they gain in the first person deliberative stance." (Bagnoli, 2021c, pp. 405–406).

Nonetheless, the defining characteristic of deliberation remains its autonomous stance and for all Kantian constructivists, as well as others,[227] autonomy has a very specific meaning: it refers to an activity of self-legislation through which certain claims become normatively binding for us (Bagnoli, 2013b, p. 120). To understand how constructivists develop a normative conception of rights, it is useful to explore how their notion of autonomous agency evolves in relation to other normative concepts, such as reasons (2.2) and values (2.3), which appear to precede the concept of rights.

[227] Normativity as self-legislation is also endorsed by non-Kantians and, though not explicitly, non-constructivists. David Velleman (Velleman, 2000a, 2009) is an interesting example in this sense, since he does not identify as a constructivist, but only as a constitutivist, and he does not present himself as a strict Kantian either, but rather, in his words, as a "Kinda Kantian." "My Kinda Kantian strategy is to argue that the aim constitutive of agency can be seen to have pushed us in the direction of our moral way of life, and to be pushing us still in directions that are recognizably moral. The strategy thus provides a retrospective commentary on morality as a rational development, a form of rational progress." (Velleman, 2009, p. 149). For Velleman, the key to understanding normativity in action is to discover what stands to action as truth stands to belief. It is constitutive of belief that it aims at the truth; truth is normative for belief, meaning that beliefs ought to be true, and a believer is subject to a requirement to be responsive to the truth. Similarly, it is constitutive of action that it aims at the good; the good is normative for action, meaning that actions ought to promote the good, and a person who acts is subject to a requirement to be responsive to the good. However, Velleman argues for the different idea that it is constitutive of agency that agents have the aim of knowing what they are doing: self-knowledge is the constitutive aim of action. This determines the criterion of success in action, which in turn determines which considerations count as reasons for action (Velleman, 2000a, pp. 368–373). It is against this background that Velleman distinguishes two senses in which the aim of self-understanding is inescapable: natural inescapability and dialectical inescapability. "[T]he aim of self-understanding is inescapable for you, and in two senses. First, it is naturally inescapable for you as a human being. As a human being, you are naturally endowed with a theoretical intelligence, which is not a passive receiver of information but an active synthesizer, striving to make incoming information hang together so as to represent an intelligible world. You are also endowed with an objective self-awareness, a concept of yourself as a part of the world to be understood. These two endowments inevitably com bine to yield the aim of understanding yourself, which inevitably motivates you to do what you can understand, whereupon you become a writer rather than a reader of your own behavior. […] Because the aim of self-understanding is naturally inescapable, the question whether to have it is moot: you cannot help having this aim, like it or not. And because you have it, you engage in a kind of reasoning by which you end up writing your forthcoming behavior, with a warrant to write it as you like, without having to read it from evidence. The aim thereby sets you to reasoning in a way that makes you autonomous, constituting you as an agent, and the resultant behavior as action. This brings me to the second sense in which the aim of self-understanding is inescapable: it is constitutively inescapable for you as an agent and hence as the kind of creature who can ask practical questions like "Why should I have this aim?" If this is indeed a practical question, posed in the spirit of deliberation between options, such as those of adopting or rejecting an aim, then it must be resolved via the sort of reasoning that is practical, in the sense that its resolution is authored rather than discovered, written rather than read. The question must therefore be resolved, I have argued, via reasoning that aims at self-understanding. To ask "why should I have the aim of making sense?" is to reveal that you already have it. If you don't seek to do what makes sense, then you are not in the business of practical reasoning, and so you cannot demand reasons for acting or aiming." (Velleman, 2009, p. 136).

2.2 REASONING ABOUT REASONS

At least since 1957, with Elizabeth Anscombe's seminal work *Intention*, the focus of the debate in the philosophy of action—and practical philosophy more broadly—has shifted from the faculty of reason to reasons: the considerations that guide or justify our actions (Alvarez & Way, 2024). Anscombe argued for a crucial distinction between intention and intentionality, stressing that objective reasons are grounded in external purposes and observable actions, rather than merely in the internal intentions of the agent.[228] This opened a multitude of different but interconnected debates, that are often grouped in the so-called agent-relative/agent-neutral distinction.[229] Michael Ridge (2023) has classified the distinctions into six families,[230] including the distinction between internalism and externalism and intersubjectivity or non-intersubjectivity.[231] Other philosophers

[228] Intention refers to a specific mental state or attitude that guides an agent's actions, involving a commitment to carrying out a particular action or achieving a certain outcome. Intentionality, on the other hand, is a broader concept that refers to the quality of mental states and actions being directed towards an object or state of affairs. Intentionality encompasses not only intentions but also other mental states that involve some form of direction or focus on an object or outcome: it is the aboutness or directedness of mental states such as beliefs, desires, hopes, and fears. Reason enters into Anscombe's distinction through its crucial role in practical deliberation, by forming and justifying intentions and influencing the directedness of various mental states (Anscombe, 1956)

[229] "The agent-relative/agent-neutral distinction is extremely important to a wide range of debates in normative philosophy. Yet the distinction is often drawn in very different ways, with the risk that philosophers are simply talking past one another" (Ridge, 2023).

[230] Each families is structured as a rival pair, to parallel the key aspects of the agent-relative/agent-neutral distinction: 1) Internal/external 2) universality/non-universality or generality/non-generality 3) deliberator-relative/deliberator-neutral or bindingness-relative/ bindingness-neutral 4) appraiser-relative/appraiser-neutral 5) intersubjectivity/non-intersubjectivity or essentially-shared/not-essentially-shared 6) reasons-for-something-to-happen/reasons-to-do.

[231] "Once intersubjectivity/non-intersubjectivity is explicitly defined, it is relatively clear that it is different from the agent-relative/agent-neutral distinction. For the former distinction is drawn in terms of communicability and the latter distinction makes no reference whatsoever to communicability. Indeed, for all that has been said so far, a reason might be intersubjective and be either agent-relative or agent-neutral. Still, it is surprisingly easy to run together these two distinctions. One way this confusion might arise would happen in two stages. First, one might confuse the agent-relative/agent-neutral distinction with the distinction between reasons that are essentially *shared* and those that are not. Since intersubjective reasons are often characterized as reasons "we can *share*," the stage is set for confusing both of those two distinctions with the further distinction between intersubjective and non-intersubjective reasons. For it would not be that difficult to confuse shareability with being essentially shared." (Ridge, 2023).

articulate the division differently, interconnecting and mixing some of these elements (Hammerton, 2019; McNaughton & Rawling, 1991).[232]

Constructivists too think that there is a lot of confusion on the distinction between agent-relative and agent-neutral (Bagnoli, 2021a; Korsgaard, 1993). In order to clarify the conception of reason they share, I will firstly tackle the internalist-externalist debate (2.2.1) trying to elucidate the vocabulary constructivist use when they speak of normative (Korsgaard, 1996b),[233] impartial (Darwall, 1983), justificatory (Forst, 2002) and probatory reasons (Bagnoli, 2013b), as distinct from motivating reasons. I will then explore in which sense the concept of public (Rawls, 1971) or intersubjective reasons (Habermas, 1985) is an alternative to subjective and objective reasons (2.2.2), what type of practical reasoning constructivists subscribes too (2.2.3) and what is the relationship between the empirical elements and the formal core in moral reasoning (2.2.4).

2.2.1 Motivation and Justification: Internalism Vs Externalism

If reasons are commonly understood as considerations in favor of an action, belief, or attitude, then this concept appears to address at least two distinct kinds of questions: one about how one ought to act (a normative reason) and another about what actually motivates someone's to act (a motivating reason). For instance, we can ask whether there is a reason for the government of a country to tax sugary drinks (normative) and ask also for the government's reason for *actually*

[232] See for example Brink's puzzle, which revolves around reconciling four intuitively plausible claims about morality and reasons for action: "I Moral requirements-including other-regarding obligations-apply to agents independently of their aims or interests. II. Moral requirements necessarily provide agents with compelling or overriding reasons for action. III. Reasons for action are dependent on the aims or interests of the agent who has them. IV There is no necessary connection between other-regarding action and any aim or interest of the agent." (Brink 1992, 22). Here all the different claims touch on many of the agent-relative agent-neutral debate, but especially the pair motivation-justification and objective-subjective.

[233] Korsgaard initially referred to agent-neutral reasons as "deontological reasons" by using Nagel's terminology (Korsgaard, 1993, pp. 28–39) but after *The Source of Normativity* she take the terminology of Williams and use mostly "normative reason" (Korsgaard, 1996b, p. 252). Nonetheless, Korsgaard sometimes seems to prefer distinguishing primarily between private and public reasons, suggesting that objective reasons can only be understood in the public sense. In replying to some critics, she wrote: "there are those who argue that a universalizability requirement, even if it is in a sense legitimate, cannot bridge the gap from what in lecture 4 I called 'private reasons' (subjective, agent-relative, or egoistic reasons) to what I called 'public reasons' (objective, agent-neutral, or altruistic reasons)" (Korsgaard, 1996b, p. 221). It is important to bear in mind that such distinctions are valuable insofar as they pertain to specific problems inside specific debates, but they risk becoming misleading if extrapolated from that context and applied to other domains.

taxing the drinks (motivating). The same reason may address both questions: the reason that supports taxing sugary drinks could be that the tax will help reduce child obesity. In such a case, the government is motivated by a reason that not only guides the action but also justifies it. However, this example shows that actions are not always driven by justifiable reasons. For example, the government might tax sugary drinks because some of its members own shares in companies that sell low-sugar alternatives. In this scenario, the motivating reason behind the government's decision is not a reason that justifies the action. Thus, the distinction between normative and motivating reasons allows us to differentiate between the psychological question of what motivates agents to act and the normative question of whether those reasons are justifiable (Alvarez & Way, 2024).

The internalism-externalism debate is often framed inside this distinction, by trying to specify the gap between the two concepts: internalism posits that there is a strong connection between normative reasons and motivation, whereas externalism maintains that normative reasons can exist independently of an individual's motivations (Finlay & Schroeder, 2017).[234] Internalists argue that a person cannot recognize a normative reason without it becoming a motivating reason if they hold the judgment genuinely; so, if someone believes they ought to help others, this recognition is thought to inherently motivate them to act (Falk, 1986). On the other hand, internalists hold that motivating and normative reasons are two distinct orders of reasons. Normative reasons provide justification for actions, regardless of whether they actually motivate the agent (Parfit, 1984).

Whitin this debate, constructivists will typically side with internalism (Bagnoli, 2013b),[235] since the majority of them recognizing that a normative reason already provides an adequate reason

[234] These terms were introduced by W. D. Falk (1948) which divide between those who regard motivation as internal and those who regard it as external to obligation. Bernard Williams reignited this debate in 1979, developing his internal reasons theory, arguing that a person has a reason to act only if that reason is connected to their existing motivations or "subjective motivational set" (Williams, 1979). Constructivists reject Williams position because, as we saw and as we will further explore, they understand reasoning a source of reason (not a detector of it), and therefore the distinction between external and internal is dynamic, not fixed and static. Contemporary philosophers like Stephen Finlay and Mark Schroeder have further refined the debate between internalism and externalism by clarifying that the core issue is not centered around motivating reasons, but if normative reasons are also necessarily connected to motivation (Finlay & Schroeder, 2017). With regard to this particular framing of the debate, constructivists will side mostly with internalists, as our discussion will reveal. See the next note.

[235] "In order to qualify as probatory or normative reasons, they must be conceived as authoritative for the rational agent" (Bagnoli, 2013b, p. 125).

for action in so far as you are rational (Korsgaard, 1986, pp. 24–25).[236] The motivation is derived directly from the reflexive nature of practical reason: the agent's need to refer to constitutive principles of reasons, which reflects a practical necessity, rather than a purely logical or ontological one. This necessity emerges in response to the question "What should I do?", engaging us as reflective and responsible (Bagnoli, 2019; Korsgaard, 2019) yet fallible (O'Neill, 1989) agents who must navigate principles of practical reason. At the conclusion of practical deliberation, the adoption of a principle for action involves the agent in a form of rational self-binding, committing to that principle both normatively and motivationally.

However, while the internalism-externalism debate focuses on whether normative reasons inherently motivate us,[237] the distinction between normative and motivating reasons has broader implications for the relationship between motivation and justification. The concept of a motivating reason, in fact, sometimes conflates two distinct types: motivating and explanatory reasons. To illustrate this distinction, Maria Alvarez provide and illuminating example using Shakespeare's *Othello* (Alvarez & Way, 2024).[238] Othello kills Desdemona because he believes she has been unfaithful, a belief fueled by the lies of Iago; Desdemona, however, is actually innocent. Here, we observe two distinct types of reasons: one that motivates Othello—the supposition of Desdemona's

[236] "Kant maintained that, if we thought about it, we would see that we are not immune to the laws of pure practical reason: that we know we can do what we ought. But there is no guarantee of this; for our knowledge of our motives is limited. The conclusion is that, if we are rational, we will act as the categorical imperative directs. But we are not necessarily rational. […] The fact that the law might not govern conduct, even when someone understood it, is no reason for skepticism: the necessity is in the law, and not in us." (Korsgaard, 1986, pp. 24–25).

[237] "It is important to clarify that reasons internalism is a thesis about *normative* (or *justifying*) reasons, not about *motivating* (or *explanatory*) reasons. A normative reason is a consideration that counts in favor of or against doing something, whereas a motivating reason is an answer to the question, 'why did she do it?'. Clearly, motivating reasons are connected to motivation; reasons internalism maintains the more interesting claim that normative reasons are also closely connected to motivation." (Finlay & Schroeder, 2017).

[238] "Othello kills Desdemona guided by the belief, induced by jealous Iago, that she has been unfaithful to him. The tragedy of the play lies in the fact that Desdemona is innocent: she loves Othello and is faithful to him. Clearly—and putting aside whether infidelity could ever favour murder—there is in *Othello* no reason that favours the murder: no normative reason for Othello to kill Desdemona, since she is not unfaithful. Still, there are two things that seem true about Othello's killing of Desdemona. One is that he is motivated to kill Desdemona by the (putative) fact that she has been unfaithful. The other is that his action can be explained by citing the fact that he believes that Desdemona has been unfaithful. So here we seem to have two quite different reasons: one that motivates Othello—the (putative) fact that Desdemona has been unfaithful; and one that explains his action—the (actual) fact that *he believes* that she has been unfaithful. We can distinguish, that is, between the reason that explains Othello's action (the fact that he believes that Desdemona has been unfaithful) and the reason, or apparent reason, that motivates him to act (Desdemona's alleged infidelity itself)." (Alvarez & Way, 2024).

infidelity—and one that explains his action—the (actual) fact that he believes she has been unfaithful. In other words, we can distinguish between the reason that explains Othello's action (his belief in Desdemona's infidelity) and the reason, or apparent reason, that motivates him to act (Desdemona's alleged unfaithfulness), even if both of these reasons are not justifications for his actions. This gives us a more nuanced framework with three distinct types of reasons corresponding to a different function: normative reasons favor or justify an action, motivating reasons drive or influence an action, and explanatory reasons clarify or account for why the action occurred.

Bagnoli similarly distinguishes between explanatory, operative and justificatory (or normative or probatory) reasons: "we deploy reasons in order to explain, evaluate, or justify what to do" (Bagnoli, 2013b, p. 112).[239] When we seek to explain an action, we may attribute some reasons for action to an agent, but attributing normative reasons from a third-person perspective may not align with the agent's own first-person perspective of those reasons. [240] This asymmetry in what should motivate an action according to external standards and what actually motivates individuals from their own perspective, determine the distinction between normative and operative reasons. Agents are not always in the best position to articulate their own actions in terms of their underlying reasons; they may fail to fully grasp which considerations actually drive their actions.

[239] Bagnoli calls normative reasons "probative reasons" in analogy with legal terminology. "The probative value of reasons is understood here in its original legal sense as "tending to prove," that is, having the quality or function of proving, demonstrating, affording evidence for something to be the case. In order to have probative value, reasons must establish the truth of the matter. This might seem to require us to endorse strict evidentialism, the view that for something to be rational it is to be constrained by how the world is. On this reading, reasons have probatory role when they track some truths in the world. When understood as cognitivist irrealism, Constructivism does not deny that reasons may have probatory role and serve as a basis for assessing evidence that establish the truth of the matter. What constructivists deny is that the probatory value of reasons is completely guided by the world, or governed by realist standards (Bagnoli, 2013b, p. 124). The terms operative reasons was instead widely used inside the internalist-externalist debate (Falk, 1986), including by constructivists (Korsgaard, 1986). Insofar as one is rational, one's normative reasons are operative, that is, exciting and motivating.

[240] "Explanatory reasons are reasons that make one's attitude or action intelligible. Suppose Boris walks out of a department meeting to take a walk in the lake park. When colleagues ask why he walked out of the meeting, Boris is not expected (only) to offer an explanation of his behaviour, but (also) a justification of an action that appears to violate departmental duties. Did he feel sick? If so, he should be excused. Was he obliged by a stronger obligation? Then, he really did not violate any obligation, because the stronger obligation cancelled or silenced any other consideration. Or was he simply taking the liberty of disregarding his duties to satisfy his desire for fresh air? Well, in this case Boris is to be blamed for making exception in his favour. Justificatory or normative reasons are considerations in support of or against actions or attitudes, which make them normatively significant. We appeal to normative reasons to justify actions or attitudes; they are considerations in favour of or against doing a certain action, harbouring a certain attitude, or holding a certain belief" (Bagnoli, 2013b, p. 112).

Therefore, it is possible—and often the case—that observers can propose more insightful explanations of why agents acted in a particular way. This occurs because the agent may be influenced by factors that are less clear to themselves than to outside observers.

Even so, agents hold a unique authority over their actions; therefore, while the distinction between operative and normative reasons aids in attributing actions, its applicability in explaining action remains debatable (Bagnoli, 2013b, p. 113). For constructivists, explaining an action requires understanding the reasons the agent personally endorses, since acting with agency means being (and feeling) guided by one's own reasons. In the Kantian version of constructivism, this authority implies autonomy, "the capacity for self-government and normative guidance" (Bagnoli, 2013b, p. 118), thus agents have a unique role in self-governing and guiding their actions based on reasons they accept as valid (Bagnoli, 2013b, p. 117).[241] By understanding autonomy as self-legislation, autonomous decisions are not only authentic and personal, but also authoritative because principled (Bagnoli, 2013b, p. 122, 2016, p. 1245). The binding force of these normative but constructed reasons comes from their alignment with rational principles that agents, in their capacity as autonomous beings, recognize as necessary or constitutive of rational action (Korsgaard, 1996b, p. 252).[242]

The crucial aspect that is important for all constructivist, not only Kantians,[243] is to specify in which sense the normative reason can justify any actions by taking distance from a foundationalist realist framework of normative independent facts and insisting on the strong link

[241] "When we raise questions about what to do we are engaging action as agents, rather than as bystanders. We are asking for considerations that we endorse as reasons for action. This is a sort of operation that is open to us only insofar as we are the prospective agents of the action, and in this specific sense we have a special claim on it. […] According to Kantian constructivism, to be the agent is to be the author of the action, and the special claim authors have on their actions is one of agential authority. Agents exercise a special authority on their actions insofar as they are theirs. Agential authority implies autonomy, the capacity for self-government and normative guidance." (Bagnoli, 2013b, p. 117).

[242] "Normative reasons, in my view, always come from reflective endorsement. The explanation provides the material with which reflection works in endorsing or rejecting an obligation, but it does not do the work itself." (Korsgaard, 1996b, p. 252).

[243] "But all forms of constructivism agree in rejecting the recognitional realism of the naive practical standpoint. Even non-Kantian versions of constructivism are committed to this aspect of Kant's "Copernican revolution" in the practical sphere. Practical normative facts are not just "there" to be recognized or not. Rather their "way of being" is within the practical standpoint, as presuppositions and commitments of practical reasoning." (Darwall, 2006, pp. 293–294)

between agency and normativity. As Darwall puts it: "What makes someone a constructivist in our current sense is simply that she holds that the existence of normative reasons depends upon what a rational agent, or what someone deliberating about what to do ration ally, would take as reasons and be motivated by. The central claim is that it is the rationality of the agent or deliberative procedure that determines normative reasons, rather than vice versa." (Darwall, 2006, p. 293). For Kantian constructivism specifically, the link between agency and normativity is autonomy as a legislating activity, and the constructions of reason claim universal authority even if normative truths are not prior and independent of practical reason, but are determined by the activity of reason itself (Bagnoli, 2013b, p. 122). Practical reason is thus an active faculty governed by specific normative criteria that are internal and constitutive of that activity. Reasons provide evaluative standards, but they do not have authority on us unless we conceive of them as laws for ourselves through self-legislation (Bagnoli, 2013b, p. 121; Korsgaard, 2008, p. 317).[244]

Nonetheless, some scholars argue against this link between agency and normativity, insisting precisely that normative reasons are, by definitions, agent-neutral.[245] Only objective reasons, agent-neutral reasons valid for every person without considerations of his or her concrete identity, are good action-justifying reasons (Nagel, 1970).[246] In order to unpack all the assumptions underlining this position, we have to see if subjective and objective reason are a pair that parallel normative and motivating reasons.

[244] "Constructivism does not dispute that reasons are facts or that moral judgments have cognitive cores; nor does it dispute that reasons have both cognitive and affective cores. It disputes that normative reasons are normative insofar as they are true facts. Facts cannot be normative reasons, unless they are self-legislated. In other words, facts become normatively significant and acquire authority on our mind only through self-legislation. This is not an ontological claim about the nature of the contents of reasons, but a claim about the source of their authority." (Bagnoli, 2013b, p. 121).

[245] Thomas Nagel writes that the "basic question of practical reason from which ethics begins is not "What shall I Do?" but "What should this person do?" (Nagel, 1986, p. 141). Notice the similarities with Forst's version: "The basic moral question is not Who do I want to be? but What can I justify morally?" (Forst, 2002, p. 256).

[246] "Whenever one acts for a reason, I maintain, it must be possible regard oneself as acting for an objective reason, and promoting an objectively valuable end" (Nagel, 1970, pp. 96–97).

2.2.2 Subjective, Objective and Intersubjective Reasons

As discussed, the concept of 'types of reasons' varies and is often specified stipulatively, allowing some constructivists, like Rainer Forst, to describe constructivism as a form of externalism: "Internalism emphasizes the 'subjective' side of these action-motivating reasons, externalism their independent, action-justifying validity. It therefore makes sense to distinguish between reasons that explain an action and ones that justify an action; between reasons a person had for doing something and reasons to justify it" (Forst, 2002, p. 248). Here, another key conceptual pair is introduced that is as recurrent as normative and motivating reasons: *subjective* and *objective* reasons.

In both common language and philosophy of action, subjective reasons often overlap with motivating reasons, while normative reasons are often equated with objective reasons. However, using these terms interchangeably risks oversimplification and overlooks some of the important distinctions we made, like explanatory reasons, and set up an overly strict dichotomy between justification and motivation. As we said, it is not a contradiction to say that a reason can be both motivating and normative (internalists argue precisely that a normative reason should also be motivating).[247] Instead, in everyday speaking, it seems a contradiction in terms of saying that a reason is at the same time subjective and objective, while it is perfectly fine to have an objective motivating reason, a valid consideration that aptly move the agent to act, or a subjective normative reason, a considerations that the agent deems relevant based on their partial understanding of the situation (what Bagnoli would define operative reasons).

This is due to the fact that most of the time subjective reasons are intended as reflecting a person's internal perspective, varying according to individual beliefs, desires, or incomplete understandings, whereas objective reasons derive from stable, external facts that are universally valid and unaffected by individual viewpoints, providing a firm basis for moral justification. This external foundation allows objective reasons to guide motivation without being reduced to purely

[247] If someone drinks water because they know they are dehydrated, the reason is both motivating (it drives action) and objectively justified (when you are dehydrated you ought to drink water). If I don't know the importance of being hydrated, but I am only moved by the phenomenological unpleasant experience of being thirsty, I would have a subjective normative reason to drink. More specifically, I would first have an unpleasant sensation (the dryness in my throat), that would form a belief (that drinking would alleviate my throat) and then I would act on that. See Scanlon example of thirstiness (Scanlon, 1998, p. 38) and the parallel with Warren Quinn's example of a man who feels an urge to turn on every radio he sees (W. Quinn, 1993, pp. 236–237).

subjective interpretations. However, framing the questions in these terms risks conflating the notion of subjective reasons with that of desire, which is a point of tension between internalist (M. Smith, 1987; Williams, 1979) and externalist (Dancy, 2002; Parfit, 1984),[248] between realists about reasons (Broome, 2001; Nagel, 1970) and constructivists about reasons (O'Neill, 1989; Rawls, 1955), but also between Humean constructivist, which provide desire-based theories that relax the unconditionality of moral reasons (Street, 2008; Wong, 2008) and Kantians constructivist which invoke rational constraints as constitutive of the activity of willing (Bagnoli, 2021c; Korsgaard, 2008).[249]

In general, it seems uncontroversial to say that desire motivates us to act, at least in the broad sense when they are just conceptualized as anything capable of moving us to act, namely any pro-attitude that an agent may have. Nagel however distinguishes between unmotivated and motivated desires by separating the one that are not dependent on some other state for their motivation and reason giving-force: an unmotivated desire is one which is simply caused in us, while a motivated desire is one for which we can give reasons (Nagel, 1970, p. 29). Yet this distinction appears merely to restate the question of how and why an unmotivated desire can be caused in us. [250] As Korsgaard puts it: "To have a personal project or ambition is not to desire a special object which you think is good for you subjectively, but rather to want to stand in a special relationship to something you think is good objectively." (Korsgaard, 1993, p. 37).

Building on Nagel's realism of reasons, numerous metaethical theories have emerged, including Scanlon's contractualism, which emphasizes that desires are only a secondary concept

[248] If internalism were true, reasons could not, according to Parfit, guide us independently of the desires we already possess. Yet, it seems clearly evident that the common meaning of reason includes independence from current desires, as shown by the notion of advice, a concept used daily even by internalists, which would lose its meaning since nothing could be advised to anyone outside the desires they already have. The concept of advice is already a normative concept and implies normativity independent of desires (Parfit, 1984). Internalists typically respond by reinterpreting the conditions under which normative reasons become motivating. They argue that if someone fully understands a normative reason yet remains unmotivated, this suggests either that they do not genuinely accept it as a reason, or that they lack the necessary desires or dispositions to act in accordance with it (M. Smith, 1987).

[249] The will, as practical reason, is precisely the connection between desire and reasons for Kant. As Stephen Engstrom brilliantly put it: "In identifying the will with practical reason, Kant reverses the order of concepts in the traditional Scholastic conception of the will as rational desire, reconceiving the will as desiderative reason." (Engstrom, 2009, p. 25)

[250] We will also examine how Forst responds to Nagel on the next pages—a response that, while similar to Korsgaard's, approaches the issue from a different perspective.

derived from reasons (Scanlon, 1998, p. 49). Constructivists instead generally only endorse the minimal, intuitive idea that not all reasons for action are desire-based: often we do things we have no desire to do (Darwall, 1983).[251] Other constructivists take an even stronger position, arguing that desire and reasons for action cannot be conceived separately but must instead be understood as expressions of a single faculty, both possessing cognitive content. Kantian constructivists, for example, endorse this thesis and further develop it by framing reasoning as self-legislation, drawing attention to the critical distinction between personal and principled decisions (Bagnoli, 2013b, p. 122; Korsgaard, 1986, p. 22).

This issue connects to the broader debate between cognitivism and non-cognitivism, where subjective aligns with *prescriptive* and objective with *descriptive* meanings. Cognitivists argue that moral statements can be true or false, while non-cognitivists contend that they express attitudes (like approval or disapproval) rather than truth. Gibbard's expressivism, a recent evolution in non-cognitivism, refines earlier emotivist ideas by shifting the focus away from simply defining moral terms like "good" or "right". Instead, it emphasizes the significance of the attitudes people hold toward these terms, [252] a shift intended to avoid issues such as the Frege-Geach problem that challenged Ayer's emotivism.[253] For Gibbard, humans express normative judgments by aligning with standards that would remain rational and coherent, grounded in an interplay of apt feelings

[251] Some add that desires, strictly speaking, lack propositional structure and are not themselves reasons for action: only propositions like "A desires x" can serve as reasons. Thus, reasons for action are not desires per se but may be at most "desire-based" (Darwall, 1983, pp. 18–28). Darwall' account is internalist (O'Neill, 1987, p. 61) because he does not need a separate theory of motivation to explain why reasons can move agents, but only specifying which types of reasons can do so. Darwall refers to them as *impartial reasons*, reasons that hold weight beyond an individual's perspective, aligning with broader standards of justification in an impartial standpoint. (Darwall, 1983, p. 42).

[252] This position might be mistaken as a stronger form of subjectivism, in a sense of a "barrier to shared understanding" that we will examine in later pages. This possible misinterpretation shows why the pair "subjective-objective" might be inadequate terms in this context if not properly specified. Gibbard does not intend to deny ethical objectivity; instead, humans make normative judgments not by referencing personal feelings but by considering that a specific norm would still hold even if I didn't endorse it. The emphasis remains on rationality, central to coherence and acceptability, while grounded in emotions and feelings.

[253] Although some scholars think that even Gibbard version does not appropriately answer the problem (Schroeder, 2008).

and appropriate reasons.[254] For example, it would be unreasonable to blame someone for merely delivering bad news, despite any strong negative feelings about the news itself.[255]

While there are similarities within expressivism (Korsgaard, 2021, p. 188),[256] constructivists mostly hold a cognitivist position.[257] Moral judgments can be true or false without aiming to represent any objective normative reality. The "truth-makers" of these statements come from within the practical perspective itself, so that descriptive-objectivity arises from standards internal to practical reasoning rather than from independent, metaphysical facts. Bagnoli in particular explicitly defend Kantian constructivism as a form of irrealist cognitivism (Bagnoli, 2013b, p. 124).[258] Normative truths can be established (contrary to what non-cognitivists believe), but they can be non-arbitrarily constructed by practical reasoning (rather than found in the world as realists maintain). Constructivism, thus, does not dispute that reasons are facts (Nagel, 1979) or

[254] In Gibbard's theory, emotions—traditionally seen as opposing reason—become central to the concept of rationality as the foundation of normative judgments. Moving beyond the emotion-reason dichotomy is also fundamental to constructivism (Bagnoli, 2011).

[255] This angry reaction toward a "messenger who bears no blame" should be condemned as unreasonable (rather than irrational), though it remains understandable when considering the agent's state of mind. Humans are rational, finite, and fragile agents, meaning that a conception of rationality is meaningful only when considered within this human finitude. In this sense, there is an interesting common ground between expressivism and constructivism (O'Neill, 1989).

[256] Some have even suggested that expressivism may offer the most appropriate semantics for constructivism, but what seems to be lacking in the expressivism position is the sense of authority provided by valid reasoning (Bagnoli, 2021c, p. 408).

[257] Still, constructivism also aims to bridge the gap between these opposing perspectives, reflecting a general trend in contemporary moral philosophy toward less polarized debates (Bagnoli, 2024). Originally, non-cognitivism emerged as a clear alternative to cognitive and realist theories, presenting moral judgments as expressions of emotions or attitudes rather than statements of fact. Over time, however, its radical claims have moderated. For instance, while early non-cognitivists claimed moral judgments lacked descriptive meaning, the view has evolved to see such meanings as secondary. Similarly, instead of denying truth-value to moral judgments, contemporary non-cognitivists often interpret truth in a minimal or deflationary sense, acknowledging limited forms of truth and falsity.

[258] "Kantian constructivism can and should be defended as a form of irrealist cognitivism, according to which moral judgments are true or false statements whose purpose is not to represent a normative reality. Objectivity is practical rather than metaphysical, but this simply means that there are no facts of the matter that make normative statements true or false. Kantian constructivism offers a distinctive account of the truth-makers of normative statement that are internal to the practical standpoint. This is why the standards of objectivity are better said practical rather than metaphysical." (Bagnoli, 2013b, p. 124).

that moral judgments have cognitive cores (Parfit, 1984); It disputes that normative reasons are normative insofar as they are true objective facts (Bagnoli, 2013b, p. 121).[259]

The terms subjective and objective reason, however, are also used by scholars in a further different sense, with subjective being synonymous with incommunicability. This framing emphasizes a third pivotal concept for constructivists, that of intersubjective (or public) reason.[260] Kant notoriously introduced this term, but it was John Rawls and Jürgen Habermas who made significant contributions to the development of this concept within modern moral and political philosophy.[261] Rawls (1971, 1993, 1997) emphasized a distinction between public reason and private or comprehensive doctrines. He argued that moral and political principles should be justifiable to others in terms that are accessible to all reasonable citizens, regardless of their personal or cultural perspectives. Public reason, according to Rawls, is the sphere in which citizens offer reasons for moral and political decisions and principles that should be justifiable to others in terms that are accessible to all reasonable citizens, not just from one's personal perspective (Rawls, 1997, p. 786).[262]

Jürgen Habermas (1985, 1996, 1998), in his theory of communicative action, further stressed the concept by positing that rationality itself emerges through intersubjective discourse. Communication, for Habermas, is not just a way of exchanging information, but the medium through which individuals engage in mutual justification and construct shared understandings: people reach valid conclusions through a process of public and reciprocal argumentation, where

[259] "Facts cannot be normative reasons, unless they are self-legislated. In other words, facts become normatively significant and acquire authority on our mind only through self-legislation. This is not an ontological claim about the nature of the contents of reasons, but a claim about the source of their authority." (Bagnoli, 2013b, p. 121). On the risk of using the notion of facts as opposed to that of values we will return in paragraph 2.3.1.

[260] Richardson prefer using openness over public, because he sees publicity has unavoidably connected to the political institution of the public, which is not central regarding to morality (Richardson, 2018, p. 169). Indeed, contemporary debates on public reason are now largely framed around the distinction between convergence and consensus approaches within liberalism (Gaus, 2010; Quong, 2011)

[261] To explore in detail the relationship and debate between Habermas and Rawls, as well as their dual connection to Kant, refer to Finlayson (2019). Although Rawls and Habermas share important points of convergence, their approaches also diverge in significant ways, since each theory of public reason is articulated in response to distinct philosophical problems. We will return in detail to both their commonalities and their divergences when examining their different conceptions of universality in Chapter 5.

[262] "Public justification is not simply valid reasoning, but argument addressed to others" (Rawls, 1997, p. 786).

all participants have the opportunity to contribute, challenge, and reason collectively. This process ensures that reasons are not merely imposed but are rather the result of deliberative engagement which is valid " just in case the foreseeable consequences and side-effects of its general observance for the interests and value-orientations of *each* individual could be *jointly* accepted by *all* concerned without coercion" (Habermas, 1998, p. 42).

These insights on public and intersubjective reasoning were inherited and further developed by contemporary constructivist scholars. Forst, for example, use the concept of intersubjective reasons to underline how agents justify actions discursively within a community, rather than relying on detached "objective" perspectives (Forst, 2002, pp. 242–245; 269–273).[263] By specifying the contexts of morality and ethics we grasp why some scholars might mix the realm of ethical and subjective reason with the objectivity of morality.[264] One example is Nagel's contention that "objective" reasons require objective values, in order to ensure that practical reasoning doesn't become limited to an agent's personal ethical views (Nagel, 1986, p. 165). However, this approach conflates deontological reasons with agent-relative ethical values, which, according to Forst, undermines their universality and obligatory force.

Forst believes that to justify universal moral obligations, Nagel would need to move beyond the subjective-objective dichotomy and instead recognize intersubjective reasons as a basis for justified action among persons. (Forst, 2002, pp. 242–243). Moral (normative) reasons are action-justifying reason that "go beyond the first person but do so intersubjectivistically and not objectivistically; hence moral justification requires that one not objective one's own perspective 'from outside' but expand it discursively toward others" (Forst, 2002, p. 245). Intersubjective reasons are thus not just contingently shared (Korsgaard, 1993), they are shared because agents can rely on impersonal procedures to sift and shift their initial, pre-reflective preferences, in order

[263] "[I]t is essential however to decide what answer a practical question demands. And this demand is an intersubjective demand: it is the others who demand an answer to the question Why are you doing this? Actors must ask themselves (and are asked) how they can justify themselves. To that extent, the faculty of judgment is indeed a subjective faculty, but judging itself is a form of intersubjective justification." (Forst, 2002, p. 272).

[264] "In a normative respect, persons are 'situated' in different contexts: they are members of different (ethical, legal, political, and moral) communities in which they are faced with practical questions that they must answer with good reasons within these communities. A theory of morality (in the larger sense) must therefore begin in intersubjective-practical contexts in order to reconstruct the different modes of validity and justification according to which 'autonomous' persons act 'rightly'." (Forst, 2002, p. 241).

to reach a "filtrate of shared communal attitudes" that forms the basis for the notion of intersubjective reason (Darwall, 1983, p. 143).

Bagnoli similarly contends that agents derive moral justification from reciprocal recognition, where reasons gain validity through their capacity to bridge personal and shared perspectives: "For the Kantian constructivist, it makes sense to account for objectivity from the standpoint of the agent's self-consciousness. The case for objectivity rests on the moral status of the feeling of respect. Respect is the subjective aspect of the ''fact of reason'', which is supposed to show that practical reason is real because we are capable of rational choice." (Bagnoli, 2015, p. 43). Respect, in this context, is not merely an emotional response but the subjective dimension of this fundamental intersubjective moral fact: it signals an agent's recognition of moral law and their responsiveness to reasons that extend beyond personal inclinations. [265]

To conclude and reconnect to the previous section, it is crucial to emphasize that, for constructivists, intersubjectivity is not simply the social expression of the (internal) first-person plural, as some scholars have suggested—particularly in discussions on rule-following.[266] Rather, intersubjectivity plays a deeper normative and justificatory role, shaping the way agents recognize, endorse, and apply normative reasons. There is a distinction between the first-person plural as experienced from within a rule and the first-person plural as used by an external observer: the former involves a shared commitment to a set of norms and practices, where the 'we' signifies active participation in a common framework of reasoning. In contrast, the latter—used by an external observer—refers to a collective from a detached perspective, acknowledging group

[265] We will return on the role that respect plays for constructivist, especially in the human rights debate, in the next chapter.

[266] See especially Kripke's skeptical interpretation of Wittgenstein—often referred to as *Kripkenstein*—which argues that meaning and rule-following are inherently tied to social practices. Kripke's argument challenges the idea that individuals can follow rules in isolation since if meaning were determined solely by private mental states, there would be no objective way to distinguish between correctly and incorrectly following a rule. From this constatation, Kripke derives his famous *skeptical paradox*: if no fact about an individual's past usage can fix meaning, then there is no determinate meaning at all—unless it is grounded in collective agreement. Thus, Kripke's *skeptical solution* suggests that rule-following is a fundamentally social activity. Language users do not follow rules based on some internal, self-standing grasp of meaning, but rather because they participate in a shared linguistic practice where correctness is defined by communal standards. For a constitutivist response to this paradox, see Haase (2012). "I want to suggest Wittgenstein rejects both, the Distributive and the Communal conception of the social; and that he has good reason to do so. What the Distributive and the Communal Conception of the social share is the idea that the relevant notion of a social practice is to be elucidated by talk about the relations between their speakers–their agreement or disagreement." (Haase, 2012, p. 252).

membership without necessarily adhering to or internalizing its guiding principles. Constructivism offers a solution by providing an account of normativity that strikes a balance between being purely social and completely detached. The constitutivist interpretation, in particular, develops a conception of the "generic we" that avoids both of these pitfalls (Haase, 2012, p. 253).[267]

Intersubjectivity, within this framework, ensures that reasons are not merely subjective inclinations or external imperatives but instead emerge from a dynamic process of collective validation still based on some constitutive principles of Reason. This process allows for the integration of individual reasoning within a broader, shared structure, thereby bridging the gap between subjective and objective reasons. It is through this public structure that norms gain legitimacy, as they are shaped and refined through the mutual recognition and deliberation of reasoning agents. However, the connection between reasoning and agency appears to extend beyond relationality, suggesting a conception of reason that is both reflexive and transformative—a conception we will now explore in the next subsection.

2.2.3 Practical Reasoning as Instrumental, Reflexive and Transformative

Another central distinction concerning reasons, which typically precedes all others, lies between *theoretical reasons* (reasons to believe) and *practical reasons* (reasons to act).[268] This distinction helps somehow to clarify the types of reasoning faculty we are referring to: theoretical reasoning involves addressing questions of prediction and explanation, while practical reasoning

[267] Kantian constructivists, while generally agreeing with much of Haase's position, often emphasize the risk of essentialism in this Aristotelian version of the argument, which relies on the concept of the "form of life" (Haase, 2012, p. 237).

[268] Of course, this distinction is quite nebulous since theoretical and practical rationality are deeply intertwined: determining the norms of practical rationality seems ultimately a theoretical matter (Dreier, 2004) and at the same time beliefs and theoretical reasoning are often influenced by practical considerations (Harman, 2004). Nonetheless, it is a distinction that can be traced back to the very origins of philosophy. Plato, for example, emphasized the prominent role of reason in both cognitive and theoretical tasks—such as discovering explanations for how things are and justifying how they ought to be—as well as in practical tasks like correctly managing an individual's life and behavior. The distinction was later made explicit and famous by Immanuel Kant, who differentiated between "pure reason" and "practical reason". As O'Neill underlined, "Kant's most basic move in seeking to explain the grounds and limits of human reason is his claim that practical uses of reason are more fundamental than theoretical uses of reason. He offers various lines of argument against the priority or independence of theoretical uses of reason. These include his claims that theoretical chains of reasoning are intrinsically incompletable, since they unavoidably lead to antinomies, as well as his particular form of explanation of the compatibility of human freedom and natural necessity."(O'Neill, 1989, p. 29).

involves addressing questions of action and justification.[269] Looking backward to events that have already taken place, theoretical reasoning asks why they have occurred; looking forward, it attempts to determine what is going to happen in the future. Practical reason, by contrast, typically considers a set of potential actions, none of which has yet been performed, and asks what one ought to do or what would be best to do. To put it another way, theoretical rationality should guide us to appropriate beliefs, telling us what we ought to believe, while practical rationality should allow us to achieve desired or appropriate goals, showing us what we ought to do or what it would be best to do to realize some intentions.[270]

Constructivists, drawing from the Kantian tradition, focuses extensively on practical reasoning,[271] a type of activity we are all familiar with, even if we may not always recognize it, as we are constantly engaged in it when making practical inferences, such as deciding which clothes to buy, what to eat for breakfast, which career to purse and so on. While these decisions can sometimes be biased or irrational or contrary to what we judge appropriate such in cases of akrasia, in all these situations we are nonetheless always practically deliberating, specifying which reasons is more appropriate in a particular context to carry out a specific action. Practical reasoning is usually precisely described as goal-directed reasoning from an agent's scope, and from some action

[269] The different ways in which scholars specify the differentiation between theoretical and practical reasoning rely on the general conception of reasoning they subscribe to. Some, such as Raz (2011, 2022), argue that reasoning is generally concerned with the formation or modification of belief, with practical reasoning specifically addressing the formation and modification of normative beliefs about action. Others tend to separate more decisively the two, and thus they say that practical reasoning can conclude directly in intentions (Broome, 2013) or directly in action (Tenenbaum, 2021). I will stick with a broad and general conception of reasoning that can be shared by the majority of constructivists, as an inferential process that takes some input and yields as output the formation or modification of the subject attitudes (Harman, 1986). In this sense, the output of practical reasoning is directed toward action, though whether it takes the form of a belief, an intention, or the action itself remains a subject of debate. Hower, to see in which sense practical reasoning as justification is not simply a practical inference, see Bagnoli (2018a, p. 2865).

[270] This does not mean that theoretical deliberation is solely concerned with facts of the world, while practical deliberation is solely concerned with values, as some scholars might suggest (Weber, 1989). Distinguishing the two branches in this manner recreates the fact-value dichotomy that constructivism aims to eliminate. In section 2.3, we will explore the role of values within a constructivist framework.

[271] "The sense in which reasoning is practical is that it delivers normative reasons, which are both independent of the subjects' particular desires and interests and binding in the first person. There is an epistemic gain in exercising practical reasoning. This is how practical subjects acquire practical knowledge of themselves as capable of acting for a reason, and this knowledge is distinctively efficacious" (Bagnoli, 2016, p. 1244). We will return to how constructivism achieves its practical commitment at the end of this chapter.

selected as a means to carry out the goal, to achieve the action[272]. This is called instrumental rationality, and instrumentalism is the doctrine according to which all practical inferences are means-end inferences; therefore, providing a practical justification in the instrumentalism framework is merely showing that a proposed end is a means to a further end.

This is a natural position at which to arrive at reflection; for what could possibly be adduced in support of a practical conclusion except its relation to a goal or desire already possessed? In fact, this form of rationality has usually been viewed as the single, unproblematic,[273] substantive norm of practical reason. Nonetheless, instrumentalism have been heavily criticized by many scholars, for example by Adorno and Horkheimer who underlined how instrumental reason, unchecked, can serve systems of power that perpetuate inequality, exploitation and alienation (Horkheimer, 2013; Horkheimer & Adorno, 2002). Forst (2023), drawing on the Frankfurt School, critiques instrumental reasoning for its limited, means-oriented perspective and its inability to justify ends in a way that is accountable to all affected. Reasoning must go beyond simply achieving means-end efficiency: it must involve justifying norms and actions through critical and reflexive discourse. As he puts it "we need to redefine reason in a self-critical, non-reductive as well as non-reifying, intersubjectivist way, that is, as more than mere instrumental or strategic rationality, to use the terms of Habermas' comprehensive effort to update the approach in his *Theory of Communicative Action*. For Frankfurt critical theory, such reduction was the main problem of modern notions of reason (especially in their capitalist form): they turned everything into a possible object for rational use." (Forst, 2023, p. 396).

[272] This line of thought can be traced back to the philosophy of David Hume, who famously asserted that "Reason is, and ought only to be the slave of the passions".

[273] But within the discussion several issues emerged on which no common answer was found. For example, if this conception of instrumental rationality represents a binding norm of practical reason, then we are open to rational criticism to the extent we fail to exhibit this kind of instrumental consistency, regardless of whether we want to comply with the principle or not. This is usually followed with the assumption that there is no room for rational criticism of peoples' ends, but only for Weberian *Zweckrationalität*: the rational determination of means to the realization of ends that are taken to be given, as a matter of human psychological fact, but not everyone is inclined to endorse this. In an influential paper titled *Deliberation is of Ends*, Aurel Kolnai floated precisely the idea that, very often, our goals are not definite enough to serve as the starting point for means-end reasoning (Kolnai, 1961). See the paragraph 2.3.3 for further discussion on final goods.

By underlining the reflexive and public feature of reasoning, Forst argues that ends are not taken as given, but they must be critically examined (Forst, 2024, p. 144),[274] and justified to those who are impacted by them (Forst, 2024, p. 256). This does not mean that instrumental reasoning should be entirely discarded, but only that instrumental reasoning is only one aspect of reasoning, not its entirety. For this reason, Forst distinguishes between rational grounding and reasonable justification (Forst, 2012, pp. 14–15).[275] Rational grounding is a monistic, means-end efficiency calculus, concerned solely with determining the most effective way to achieve a given goal. In contrast, reasonable justification is both reflexive and intersubjective: it involves putting yourself in a critical self-discussion regarding the end you are choosing and at the same time opening to a "public space" where your reasons can be evaluated and potentially challenged by others. This dual focus on self-reflection and intersubjective accountability expands the scope of reasoning beyond mere instrumental efficiency.[276]

[274] "Any constructivist view that entails basic moral criteria of the justification of the constructions as well as of the standing of the constructing agents must be based on a reflexive notion of the moral right to justification among equal normative authorities" (Forst, 2024, p. 144).

[275] "Regarding practical contexts, it is essential to distinguish between rational grounding (rationale Begründung) and reasonable justification (vernünftige Rechtfertigung). In all such contexts, it is a matter of answering the question 'What should I do?' (in a specific way each time); but this question does not always need to be answered with normative reasons in a demanding, or moral, sense. A rationally grounded reply to this question consists of a person considering what the appropriate means are for realizing a subjectively given end and acting according to the practical conclusion that follows from relating those ends or goals to potential means. In this sense, an intentional (and 'rationalized') action can be characterized as rationally grounded if the practical deliberation leading up to it is oriented toward specific rules that refer both to the consideration of ends and to the choice of means. According to this conception of rational grounding, action can only be 'relatively' grounded: the choice of means relative to the given aims, and the evaluation and prioritizing of ends relative to that which lies in a person's 'enlightened self-interest.' Such reasons are also part of the public game of reasoning, but only insofar as the reasons that distinguish an action as rational can potentially be comprehended by other rational beings; they do not, however, require others' acceptance for their validity as good reasons. In this context, to give a reason means, first of all, to be able to explain an action; it does not mean to be able to justify it intersubjectively. The latter level is only reached if the former rationale is challenged, and not on the basis of whether the action was rational or the most rational, but on the basis of whether it was justified or could be accounted for in an ethical or moral sense. Answering this question requires a rational capacity that makes it possible to enter the public space of normative justifications." (Forst, 2012, pp. 14–15). This distinction between rational and reasonable should not be confused with objective and inter-subjective, as we highlighted in the previous paragraph.

[276] "If the question "What should I do?" -or the question "Why did you believe you were required to do this, or that it was permissible?" -is posed in a normative, for example, moral, context, it calls for a justification that differs from the above understanding of grounding. For in seeking out morally grounded answers to those questions, it is essential to be able to provide reasons that can justify actions according to criteria that are valid within a moral context. And these validity criteria are not subject-relative in the way envisioned by the conception of practical rationality or rational grounding above: here, the point of justifying action is not to realize one's own ends and goals as rationally as possible, nor primarily to rationally assess and order one's ends; rather, what is called for here is a form of reasoning

However, the question of how we reflexively choose our ends reveals a deeper issue with instrumentalism. The problem lies not only in the dehumanizing risks of allowing instrumental reasoning to dominate all aspects of thought but also in its inherent self-contradiction, as demonstrated by Elijah Millgram (2005). Since any chain of practical instrumental inference will terminate in a desire that one just has, primitive, we are left with the problem of arbitrary desires.[277] The explanation for taking the perception of some secondary quality to provide a reason for action will involve a non-instrumental pattern of practical inference, and this is why instrumentalism is an insupportable view of practical reasoning: there are desires that cannot be justified in instrumentalist terms, and justifying it requires abandoning instrumentalism. Constructivism offers a conception of rationality which is alternative to instrumentalism precisely because it offers a way for deliberating about final ends (Richardson, 1994). Kantian constructivism, specifically, does so by introducing the categorical imperative as an overarching principle, supplementing the principle of utility, which governs only hypothetical imperatives.[278] By denying the normativity of the instrumental principle taken in isolation of non-instrumental principles, Kantian constructivists recognize the categorical imperative as the structural one (Herman, 1993, pp. 154–155; Korsgaard, 2008, pp. 27–68; O'Neill, 1989, pp. 3–4).

Korsgaard's critique of Nagel is precisely that he fails to recognize that his arguments implicitly rely on a notion of instrumental rationality, limiting the scope of his argument and preventing a more thorough understanding of moral reasoning (Korsgaard, 1993, p. 50).[279] In

---

that submits both the ends of action and the means to justification before others as those morally affected." (Forst, 2012, p. 15).

[277] Some instrumentalist thinkers have no problem in admitting that since there are no straightforward criteria for reasoning successfully on arbitrary desires, practical deliberation about final ends is not a true form of reasoning. But how is one supposed to clarify one's largest and most important ends, if not by reasoning about them in some way? Practical reasoning must consist not only in figuring out how to get what you want, but in figuring out what exactly it is you want in the first place, what is the right thing to want and to do in that place, and this is a matter of further specifying your ends. This is even more evident in applied fields, such as bioethics and especially clinical ethics. Before searching for the means to return the patient to health, the physician must first decide what health, in these circumstances, would be (Viafora, 1999).

[278] "The requirement of logical consistency may discourage an individual from pursuing self-defeating ends, but it would not be of any help in selecting which proposal of action should be pursued by a group of people. The normative pressure toward reaching a consensus is not solely a logical pressure, and while logical consistency is a necessary constraint on any proposal for action, it is insufficient to provide any shared basis for joint action." (Bagnoli, 2021c, p. 409).

[279] "In both The Possibility of Altruism and The View from Nowhere, Nagel's arguments take an unexpected turn. In both he starts from recognizably Kantian ideas, working in The Possibility of Altruism with motivation derived

general, what distinguishes Nagel, but also Scanlon (1998, pp. 1, 76) from the constructivist approach is their view that reasons are conceptually prior to Reason and reasoning. In contrast, constructivists hold that Reason and reasoning take precedence precisely by constructing the reasons (Darwall, 2006, p. 293).[280] As we have noted, Kantian constructivists further specify that reason possesses both first-person singular authority and first-person plural bindingness. Although normative truths are not prior to or independent of practical reason, they are determined by the activity of reason itself. Indeed, this is precisely how facts become normative truths and acquire authority for us—through the processes of self-legislation and public deliberation.

At its core, Kantian constructivism has, from the outset (O'Neill, 1989, p. ix), emphasized that although normativity can be attributed to instrumental reasoning (Bagnoli, 2016, p. 1237),[281] this normativity ultimately derives from the normativity of non-instrumental reasoning, particularly emphasizing the crucial role of reason's reflexive self-presentation (Bagnoli, 2016, p. 1238).[282] Furthermore, by revisiting the intersubjective dimension of reasoning, Bagnoli underscores that reasoning is transformative: engaging with others expands and reshapes one's

---

from a metaphysical conception of the person, and in The View from Nowhere with a two-standpoints account. And yet in both he ends up having to construct elaborate arguments to fend off the conclusion that his ideas will lead to utilitarianism. Why does this happen? It happens because Nagel presupposes that the business of morality is to bring something about." (Korsgaard, 1993, p. 50). Marcia Baron, building on this and also on Korsgaard arguments for the value of humanity, has presented an interpretation of Kantianism that is built on the idea that persons have a value that is properly acknowledged by honoring it rather than promoting it (Baron, 2008). We will return to this point in the section regarding the moral value (2.3.2).

[280] "The central [constructivist] claim is that it is the rationality of the agent or deliberative procedure that determines normative reasons, rather than vice versa." (Darwall, 2006, p. 293).

[281] "What I propose describes practical subjects as capable of practical concerns, but it does not make rational agency coincide with moral agency, hence it does not resolve by fiat the contrast between instrumental and non-instrumental reasons. It allows for further discussion about the authority of instrumental and non instrumental reasons. The view I advocate is minimalist in the following respects: first, it is practical rather than metaphysical; second, it reserves the title "practical subjects" for rational agents capable of self-reflection, but it does not define practical reasoning in instrumentalist or non-instrumentalist terms and admits of degrees of rational agency" (Bagnoli, 2016, p. 1237).

[282] "In debates about rational choice, it is widely agreed that the first-personal perspective is the deliberative perspective of deliberators deploying at least instrumental norms of reasoning. I am reluctant to say that agents use norms of instrumental reasoning insofar as they appear to exhibit rudimentary rationality in action. […] To explain the varieties of first personal phenomena associated with practical reasoning a capacity for self-reflection needs to be in place. Most likely, even this capacity admits of degrees. […] But to qualify as a practical subject one needs the capacity to engage in forms of practical reasoning that center on self-representation. It is plausible to reserve the capacity for practical reasoning only to practical subjects, and recognize different varieties of rational agency and associated first-personal phenomena." (Bagnoli, 2016, p. 1238).

understanding of reasons. The reasons constructed through this process may differ substantially from initial personal interests, but they acquire greater authority as a result. Practical reasoning, therefore, alters an individual's desires and concerns because it is "independent of what reasoning is *about*, and of what reasoners happen to desire or value, but not independent of *who* they are." (Bagnoli, 2016, p. 1230).

Reasoning, in this sense, is not merely deductive or based on fixed first principles. Instead, it is principled yet generative and transformative, creating new normative reasons through reflective and intersubjective activity. This is what makes it irreducible to mere instrumentalism. As Bagnoli remark: "On the constructivist view, then, the effects of practical reasoning are quite remarkable and cannot be captured solely in terms of performing, bringing about something, realizing an end or producing a string of consequences. On this rich understanding, construction names a mode of reasoning genuinely alternative to classical models". (Bagnoli, 2016, p. 1245). However, beyond the structure of practical reasoning, one crucial question remains: how does reasoning practically guide agents to construct normative reasons? This will be the final issue to address regarding reasons before we turn to the role of values within the constructivist framework.

2.2.4 Moral Reasoning between Universalizability and Defeasibility

If Practical reasoning is reasoning aimed at decision or action, and if the logic of action and choice does not vary with the subject matter, then moral reasoning is simply a subgroup of practical reasoning, aimed at what do we have to do morally, namely finding moral reason for action. Yet, philosophers disagree on the nature of moral reasons, depending on the framework they are subscribing to.[283] As noted in the introduction, Kantian constructivists emphasize that what defines morality is its universality, and the principle of universalizability is what should guide moral agents in order to find principles for acting morally; therefore, a reason for action is considered moral if it can be universalized—that is, if it is one that everyone could adopt in similar circumstances. However, some critics argue that this approach reverses the natural order of

[283] Dancy for example insists on the holistic nature of these reasons—where the relevance of a reason depends on the context—arguing that universal principles are less applicable in moral reasoning than in areas like mathematics (Dancy, 2004, p. 132). Kantian constructivists, as we will see, believes that this perspective is controversial because it misunderstands how exceptions should be principled in order to avoid arbitrariness.

practical reasoning. In daily life, moral deliberation often begins by evaluating empirical evidence for and against a particular course of action, rather than abstractly and solipsistically asking whether an action can be consistently applied universally without contradiction. Consider Scanlon example in which failing to help a person whose car has broken down would be "a serious wrong in a place where someone who is stranded overnight is likely to freeze to death, but not a serious wrong in a safe country with a mild climate." (Scanlon, 1998, p. 329).[284]

This example and many more are usually tackled inside a more general debate on the nature of reasons and the possibility of defeaters. In practical reasoning, defeaters refer to factors that undermine or alter the justification for a decision or action, even if that decision was initially well-supported. An important distinction (Koons, 2022) is between undercutting and rebutting defeaters: undercutting defeaters remove evidential support for a belief (new information reveals that the reasoning or assumptions that supported a decision are flawed) while rebutting defeaters provide evidential support for the opposite thesis of the belief (new evidence provides a stronger, more compelling reason to act differently).[285]

Constructivism generally agrees that defeasibility and "calibrating" principles to the specific circumstances of an action seems a crucial consideration for any efficient theory of reasons (Bagnoli, 2018a, p. 2860), but Kantian constructivism especially maintain that exceptions are, nevertheless, principled: any substantive moral reason is in principle defeasible, what does not change is the form of moral law, which is universalization (Bagnoli, 2018a, p. 2861; O'Neill, 1989, p. 83). Universalization is the form that best coheres with the profile of the practical subjects who engage in practical reasoning. The key idea to keep in mind is that reasons function properly only when they are universally authoritative. This is why reasons, in themselves, can be undercut or

[284] "Any plausible moral view would allow for the fact that actions that are right in one place can be wrong in another place, where people have different expectations, or where different conditions obtain […] A view that allows for such variations in what is right, by applying a fixed set of substantives moral principles to varying circumstances, is not relativism but rather what I will call 'parametric universalism'." (Scanlon, 1998, p. 329). In the following pages, we will examine Korsgaard's similar notion of provisional universality, which, however, introduces a distinct set of challenges.

[285] For example, if you reason that you should buy a particular product based on its positive reviews, but later discover that the reviews are fake or biased, the justification for your action is *undercut*. You also might decide to take a particular job based on its salary and benefits. However, if you later learn that the work environment is toxic or the job would negatively impact your health, this new information overrides the initial reasons and the justification for your action is *rebutted.*

rebutted if it is shown that another reason prevails. However, universality, being constitutive of reasoning, cannot admit defeaters.

This, of course, may seem like an attempt to have the bread buttered on both sides. Kantian Constructivism has faced such critiques by many scholars,[286] including those raised by Millgram (2005, pp. 89–121), who contends that it is overly reliant on exceptionless abstract principles of reason.[287] This, he argues, risks neglecting the contingent and contextual factors that are essential for meaningful moral reasoning. To engage with this critique, this final subparagraph will reconstruct Millgram's argument and examine the most effective responses to his concerns.

As we saw, Kantian Constructivists think that there is a structural principle which governs all activities of reason, the Categorical Imperative (CI).[288] Milgram focuses specifically on the procedural reading of the CI (CI-Procedure) which aims to validate practical reasoning by developing a procedural test for the permissibility of a proposed action (2005, pp. 89–121). According to Millgram,[289] this test can be schematized in four steps:

1) First, Identify the maxim of the action. Maxims can be codified as having a logical form, like: In circumstances C, do A, because of P.

2) Universalize the maxim and verify the theoretical commitment. This requires imagining a "perturbed social world" in which everyone in your circumstances does what you are proposing to do. If you can't imagine such a world because it is literally inconceivable or

[286] Even Humean constructivists such as Sharon Street (2008) and Julia Driver (2007) usually argue that Kantian constructivism fails to make room for contingency.

[287] Millgram categorizes the different conceptions on the universalizability of reasons as follow: "Kant's moral theory, together with its underlying view of practical reasoning, occupies the right-wing extreme on a spectrum of views about the universalizability of reasons: if it's a reason, it's always a reason. Further to the left, we have Aristotelian theories of practical reasoning, which are focused on defeasibility management. On these sorts of views, if it's a reason, it's normally or presumptively a reason on other occasions; but there (also normally) indefinitely many exceptions to the general rule, because your presumptive reason can be overridden by other considerations." (Millgram, 2005, p. 18). What I am interested to contest here is precisely that Kantian constructivism does not think that what is a reason is always a reason.

[288] O'Neill (1989) argues that the categorical imperative is the supreme principle of reason in general. Korsgaard rephrases it as the supreme principle of normativity in general (Korsgaard, 1996b, p. 104).

[289] This reconstruction of the argument is mostly based on O'Neill (1989), a perspective which is also widely supported by Korsgaard and Hermann (see also the next note).

if the intention expressed in the maxim is bound to be frustrated in such a world, then your maxim fails the contradiction in the conception test.

3) Verify the practical commitment. Ask whether there are intentions that you are bound to have simply in virtue of being a human agent, but that cannot be successfully executed if your maxim is universalized. If there are, your maxim fails the contradiction in the will test.

4) Finally, if your maxim passes both the tests, you may perform the action. If it does not, acting on your maxim is prohibited.

Among the classical examples, there is the maxim that permits lying and break promise, which fails the contradiction in conception test, [290] and the maxim advocating violence, which passes the contradiction in conception test but fails the contradiction in the will test. Let's quickly examine how this second case fits inside the CI-Procedure:

1) Maxim: To use violent means and kill when it serves your interests and goals.

2) Being a victim of violence, and being killed in particular, disrupts one's ability to act and pursue projects. If everyone acted according to this maxim, your agency would be rendered ineffective, as your plans could always be interrupted by others. However, although this

[290] If lying in given circumstances was a practice, everyone would know that it was, and they would adjust their behavior in the light of that knowledge, making acting on your maxim impossible. However, in some Kantians the distinction between the two tests regarding these maxims is less clear. Consider for example this passage from Barbara Herman: "The maxim of deceitful promising cannot be willed a universal law because the deception requires the background institution (or set of conventions) of non-deceitful promising if it is to succeed in producing belief. Universal deceitful promising annuls the institution and, in making deceitful promising impossible, reveals nonuniversal deceitful promising as a condition of the deceitful-promising maxim. This is the formal wrong-making characteristic of the maxim. The basis of this formal result is a fact about human rational agency: we depend on convention-controlled communication for information essential to normal action and choice. What the rejection of the maxim of deceit signifies is the inappropriateness of principles that exploit the vulnerability of human agents to manipulative control. Such principles fail to respect the integrity and separateness of will that is constitutive of rational agency. The CI procedure has then shown that the pursuit of our interests does not justify interfering with the integrity of another's will." (Herman, 1993, pp. 154–155).

scenario undermines personal agency, it does not violate the contradiction in conception test (a world like this remains conceivable).

3) However, the fact that your own plans and agency could be disrupted by violence is something that cannot be coherently willed. You cannot consistently will a world where your own capacity for action is perpetually at risk.

4) Therefore, the violence-endorsing maxim must be rejected, as it fails the contradiction in the will test.

This procedural scheme set out in this way has several limitations and according to Elijah Millgram is even self-refuting, since the maxim to adopt the CI-procedure itself gives rise to a contradiction in the will (Millgram, 2005, p. 95). In fact, it seems that you need a maxim to motivate anyone to act accordingly to the CI-procedure, something like the following: "When I am making up my mind what to do, I will act only on maxims that pass the CI-procedure, so as to make morally or rationally permissible decisions". Millgram calls this the CI-Maxim and argues that since successful agency requires exception (given that decisions are made without all the relevant information on sight) and if successful agency cannot be successfully executed if this maxim is universalized, the CI-maxim fails the contradiction in the will test and therefore the test of the CI-procedure itself blocks the CI-maxim. Millgram's critique of the New Kantian CI-procedure thus has the form of a reductio ad absurdum (Piper, 2012, p. 210).

There have been numerous replies to this critique. Korsgaard has tried to dodge these difficulties by claiming that constructivism is provisionally universal, which admits exception, but still aspire to absolute universality throughout revising norms (Korsgaard, 1999, p. 25).[291] Millgram however think that this type of provisional universalism still generates a self-contradicting "Revision Maxim" and that "institutionalizing exceptions of that kind amounts to

[291] "There is not reason to suppose that we can think everything in advance. When we adopt a maxim as universal law, we know that there might be cases we had not though of, which would show us that it is not universal after all. In that case, we can allow for exceptions. But so long as the commitment to revise in the face of exceptions is in place, the maxim is not merely general. It is provisionally universal". (Korsgaard, 1999, p. 25).

what is, by Kantian lights, a self-frustrating plan of action" (Millgram, 2005, p. 105).[292] The reductio ad absurdum still remains in provisionally universalism, according to Millgram.

Even if it is not properly a reductio ad absurdum, the strategy of institutionalizing exceptions seems to indeed produce perverse results, insofar as it raises the complexity of the principles of action. Complex principles are not only hard to formulate, but also difficult to learn and thus their efficacy is doubtful and this was also noted in general by other notable scholars (Hare, 1963a; Nozick, 1994). This is why Aristotelian constructivists (LeBar, 2013)[293] or even Kantian constructivists (Bagnoli, 2018a, p. 2867)[294] criticize Korsgaard's solution. In the end, the distinction between provisionally and strictly universal principles is not marginal as Korsgaard believes, since it is very costly for agents and removes the classical Kantian distinction between laws as objective principles of action and maxims as subjective principles of action.[295]

Another Kantian approach to addressing Millgram's critique is to understand the CI-procedure not merely as a test of consistency but as an expression of respect for persons.[296] This interpretation implements the focus of procedural universality with the practical effort of treating individuals as ends in themselves, aligning with Kant's broader moral philosophy. While this may seem to diverge from the Neo-Kantian procedural approach that Millgram describes, and risks, in his view, losing what he sees as the most compelling aspect of Kant's theory (Millgram, 2005, p.

[292] "When I run into a case that my (first-order) maxim does not handle properly, I will revise my (first-order) maxim to incorporate the exception, so as to be able to act on rules that I understand to be "provisionally universal" (Millgram, 2005, p. 105).

[293] Lebar thinks that Korsgaard's argument seems to lead to undesirable conclusions that we must do without principles because they are too abstract to be useful in identifying moral reasons and their defeaters, leading us "to be particularists as well" (LeBar, 2013, p. 200).

[294] "By modifying the content of the principles so as to accommodate relevant exceptions, we end up with entities too complex to be manageable and practical. Furthermore, it is likely that such sophisticated principles would appear to have little moral authority, since agents would find them difficult to learn and to be guided by." (Bagnoli, 2018a, p. 2867).

[295] "The fact that one cannot know on which principle one is acting upon is a serious moral problem. In fact, it is precisely the problem that Kantian theory is designed to identify by drawing the line between merely fortunate and morally valuable actions. Korsgaard's proposal threatens to lose what is distinctive to the Kantian approach to principles, that is, the claim that they are the very source of normative authority." (Bagnoli, 2018a, p. 2869).

[296] We will return on how constructivist specify the concept of respect in the next chapter.

111),[297] it finds strong exegetical support among many Kantian scholars, including constructivists, who emphasize the reflexive and transformative dimensions of reason that we have previously discussed (Bagnoli, 2016; Forst, 2024).

Further, even Barbara Herman seems to focus on the respect element of the categorical imperative in her book *The Practice of Moral Judgment* when she explains how moral reasoning requires us to interpret situations to determine the appropriate course of action (Herman, 1993, pp. 153–154).[298] Herman introduces the concept of "moral salience"—the idea that certain features of a situation stand out or become morally significant and noticeable to the agent—to addresses a common critique of Kantian ethics: the problem of 'relevant descriptions', which arises because any action can be described in multiple ways, leading to ambiguity about which description should guide moral deliberation (Herman, 1993, pp. 157–163). According to Herman, moral salience helps resolve this issue by orienting agents to the relevant aspects of a situation, grounded in their moral development and experience.

However, this reliance on the agent's psychological development introduces a tension with Kantian constructivism. While Herman views the CI-procedure as psychologically tied to the agent's capacity to identify moral patterns in action thanks to moral development and experience, constructivists argue that moral reasoning must remain grounded in rationality rather than empirical psychological criteria: agents in similar contexts may generate different reasons for action without failing the standards of reasoning. While moral sensibility is important, it cannot

[297] "Kantian theory is deep and interesting only when it is taken seriously."(Millgram, 2005, p. 111). However, Millgram himself later acknowledges that alternative interpretations of the CI-procedure, which consider the Kantian framework as a whole, are also possible. "Showing that one of the ideas in the Kantian portfolio is unworkable leaves a valuable and still-diversified grouping of philosophical assets. (And in fact, some New Kantians have, over the past decades, come to rest more weight on the so-called Formula of Humanity in particular)." (Millgram, 2005, p. 113).

[298] In it, she posits that "Kant makes use of a universalization procedure because good willing - the condition of all justified action and choice - is found only in maxims that have the form of universal lawgiving. A maxim is not a possible maxim of a good will, and is therefore impermissible, if its goodness (the condition of its subjective choice-worthiness) depends on its nonuniversality. This is a requirement of practical rationality, and it sets the argument of the CI procedure […] [T]he value of rational agency is to be expressed in the commitment to refrain from adopting principles that are not possible for all others of one's (rational) kind. Positively, each must view her maxims as candidates for principles that could constitute a community of free and equal persons. That is why universalization matters. Those features of a maxim that depend on the maxim's not having the form of universal lawgiving explain its impermissibility; they also point to the aspect of rational agency that the resulting deliberative presumption requires us to respect "(Herman, 1993, pp. 153–154).

replace the role of reason in practical deliberation and the capacity of judgment does not need to be explained psychologically (Bagnoli, 2018a, p. 2860 fn 40).[299]

We are led again to the question of motivation and its relationship to justification: for the constructivist, the primary motivation for an agent to adhere to the categorical imperative does not arise from yet another specified maxim. Instead, it lies in the practical condition that, as self-reflective rational beings, we are compelled to continually question what constitutes the right course of action before proceeding.[300] This feature is ingrained in our necessity of maintaining our unity as agents. This is the core of the constitutivist reading of Kantian Constructivism that Millgram have forgotten when analyzing Korsgaard; an action that violates the CI entails the disintegration of the self, that is, the loss or, rather, the failure to achieve metaphysical and practical integrity. "The work of practical deliberation is reunification, reconstitution: and the function of the principles that govern deliberation, the principles of practical reason, is self-unification." (Korsgaard, 2009, p. 126).

As we engage in this process of reasoning to maintain our integrity as agents, the only resolution that proves satisfying, the conclusion that halts the practical inquiries, is that we should act in a manner that is universalizable for all other agents as well. This is what Bagnoli calls "dialogical interpretation of the forensic analogy": "In order to be authoritative, judgments must be justified to a universal audience. To satisfy this requirement of publicity, we are not required to transcend our nature as finite rational agents. Instead, we are required to think of ourselves as standing in a personal relation of mutual respect and recognition with others and aiming to construct reasons that other rational agents could understand and, in principle, share" (Bagnoli, 2017, p. 365). Ultimately, it is the availability of reasons that transform the disparate threads of habits, desires, and inclinations into the articulated unified structure of our practical identity.

[299] "Herman's solution to the problem does not seem to be congruent with the Kantian theory of practical reasoning, insofar as it is grounded in empirical psychology. The criteria of salience are a dubious basis for moral consensus and may undermine the autonomy of reason. For this very reason, Herman's account of practical reasoning does not qualify as constructivist" (Bagnoli, 2018a, p. 2860 fn 40).

[300] Animals are primarily guided by instinct, particularly the principle of self-preservation (Korsgaard, 1996b, 2004). However, since persons can adopt a reflective stance towards all of their principles and thus interrogate the innate principle of self-maintenance, they require a different grounding principle. Kantian constructivism offers a solution by emphasizing that individuals must establish themselves as the effective and autonomous architects of their actions. This is the new grounding principle of self-constitution.

Reasons thus serve the function of articulating a perspective from which we must justify what to do. [301]

The reason why some Kantian constructivists are able to address Millgram's query in a more apt why, relies not only on the fact that they give more space to the formula of humanity in the CI-procedure, but that they take Kant philosophy as a whole, without risks reducing the capacity for reasoning to a solipsistic endeavor.[302] Already in the *Groundwork*, the third formulation of the Categorical Imperative, the "Formula of Autonomy", moves beyond a merely individualistic interpretation (G 4:432). Here Kant emphasizes not only our obligation to follow universal laws (as in the first formulation), but our very status as "legislators" of universal laws: to act morally is to act from principles that we could recognize ourselves as having authored in common with other rational beings. Our freedom and worth consist in our being sources of the very moral law that binds us (G 4:433).[303]

It is on this public foundation that Kant later builds, in *Religion within the Bounds of Bare Reason* (1793/1794), a more explicit practical application of the Moral Law. Rather than employing the CI merely as a procedure for "testing" maxims, Kant turns to the establishment of social institutions—what he calls "churches"—organized communities dedicated to promoting and

[301] "When we face a moral conflict and search for a resolution, we are not simply interested in acting to get out of a quagmire; we are interested in understanding our predicament and responding to it appropriately and responsibly, that is, on the basis of good reasons. It is important to us not that we decide something, but that we choose the action that is appropriate and expressive of the persons we are, given our appreciation of the case. Actions are not the only activities through which we express and exercise our agency: attitudes should also be counted as ways in which we make an impact on the world and touch others. Thus, an important practical task of ethical theory is to offer adequate normative criteria for understanding and guiding the agent's activities in the aftermath of moral choice." (Bagnoli, 2005, p. 169).

[302] Even though, as previously discussed, some Kantian constructivists, as in the case of Korsgaard, occasionally face similar criticisms—albeit with certain qualifications (Bagnoli, 2021c, p. 406).

[303] In recent decades, Kantians have even spoken of a fourth, implicit formulation of the Categorical Imperative: the "Kingdom of Ends Formula" (Johnson & Cureton, 2024). This formulation underscores the inherently social dimension of Kantian morality, portraying rational beings not as solitary legislators but as members of a systematic union governed by common laws. Kant expresses this idea in the injunction to "act in accordance with the maxims of a member giving universal laws for a merely possible kingdom of ends" (G 4:439); within such a kingdom, every rational being has equal standing as a co-legislator and must always be treated as an end in itself. While directly connected to the Formula of Autonomy, the Kingdom of Ends Formula also integrates the insights of the other formulations. As Kant repeatedly stresses, these are not distinct principle but multiple expressions of one and the same law: each formula "follows from" the others (G 4:431), and the concepts grounding one formulation "lead to a closely connected" concept grounding another (G 4:433).

institutionalizing moral conduct within a social framework (R 6:101).[304] Kant envisions such organizations precisely as institutions founded solely on moral principles, stripped of historical contingencies and dogmatic constraints, with the overarching aim of creating a moral commonwealth (R 6:102).[305] In this sense, Kant's churches illustrate how the formal universality of the moral law can take concrete shape in human communities: they institutionalize moral principles, provide mechanisms for moral education and self-improvement, and maintain the integrity of moral practice across generations. They are, in effect, the macro-level analog of the individual moral legislator, translating the autonomy of rational wills into a public, organized, and socially effective moral order that helps cultivating, sustaining, and guiding the shared realization of moral ends.

Let us briefly summarize what we have discussed and reconnect it to the broader inquiry into how reasoning and reasons are understood within a Kantian constructivist framework. Translating the specific contingent elements at hand in terms of rationality implies different step: the guiding formal idea of finding principles that could apply to everyone (principle of universalizability) is the starting point to detect precisely what to look for,[306] while the empirical inputs that enter the procedure of universalization are the circumstances of action (what, when, why in terms of motives, etc.).[307] Empirical inputs are thus crucial for our interaction with the

[304] I am grateful to Stephen Palmquist (2022), who, in his review of the preprint of my article (Morelli, 2023), encouraged me to examine more closely how Kantian "churches" serve to promote and institutionalize moral behavior within a social framework. For a detailed constructivist reading on the role of Kantian churches and shared agency, see Bagnoli (2025).

[305] Kant enumerates four defining features of such institutions: 1) Universality and Unity: the church must rest on principles that necessarily lead to universal union, avoiding sectarian divisions grounded in accidental opinions; 2) Purity: the community must be held together solely by moral incentives, purified from the "folly of superstition" and the "madness of enthusiasm"; 3) Freedom: relationships among members, and between the church and political authority, must be based on freedom; no coercive hierarchy or arbitrary claims to private illumination may prevail; 4) Immutability: its constitution must be grounded in secure a priori principles, "as in a book of laws," rather than in mutable and contradictory arbitrary creeds.

[306] "Moral deliberation is called for by circumstances of moral conflict (or uncertainty) that are flagged by the deliberative principles (presumptions) derived from the CI procedure. If they are to be able to resolve complex and unfamiliar cases, agents must not only recognize the morally significant features of their circumstances and of their proposed action, but they must also be able to make the necessary value translation of the flagged features of action and circumstance into the terms of rational agency. The point of the value translation is to permit deliberation to take place in its terms. What a moral agent should do is to "respect rational agency"" (Herman, 1993, p. 157).

[307] "The overriding or overridden status of claims is determined by reasoning over the circumstances of action. The point is that the normative status of moral reasons is not fixed by the material characterization of their moral content." (Bagnoli, 2016, p. 1243)Many elements, both circumstantial and constitutive, come into play in specific acts of reasoning, but they are not purely 'empirical' in the sense that they belong to moral psychology and

worlds (they are the material of construction) but in order to translating the specific contingent elements into rational terms, we must begin with the guiding formal idea of finding principles that could apply universally (this is the constitutive principle of construction). While empirical inputs play a role, they are not the starting point and the criteria of correctness of reasoning are not determined by the conformity with them.[308] Moral reasons emerge as the end result of addressing a specific problem, where the circumstances of actions are filtered through the procedure of universalization (reasons are the result of the construction).[309]

The reasons we obtain at the end are intersubjective (Forst, 2002, pp. 242–245; 269–273), because the process of constructions is not solipsistically developed as Milligram reconstruct it; instead, it is a collective and public process, not only on the guiding formal idea of universalizability but also in the practical task of retrieving empirical information and interacting with other agents. This intersubjective dimension is essential: it situates reason-giving within the practical conditions of human coexistence and underscores that justification is also social practice, not merely a cognitive achievement

This interpretation effectively captures how Kant acknowledges the embodied nature of human agency, along with its temporal and situated characteristics (Bagnoli, 2022b). Both empirical and philosophical-anthropological factors play a role here, but the overarching aim remains practical. Kantian constructivism, in this interpretation, highlights the need to confront contingency and finitude but also proposes to respond to this need by offering a model of practical reasoning suitable for finite rational beings, rather than a theory of finite rationality. The difference is substantial: not beings with finite rationality, but finite rational beings (O'Neill, 1989, p. 74).

---

the minimal profile of agency. It's always necessary to distinguish between the Moral law as the form of CI and CIs as the result of the reasoning of particular rational agents. P's particular reasoning is "I want to lie to Francesco to get X, can I do it?" and the answer is "You must not lie to Francesco"; this is a particular categorical imperative, the form of which is the universal moral law.

[308] "Constructivism purports to provide an account of rational justification that is congruent with ordinary practices, even though such practices do not dictate the criteria of correctness." (Bagnoli, 2016, p. 1230)

[309] In this respect moral reasons "have no distinct normative authority and no special normative status in respect to other sorts of normative reasons. […] Moral reasons may override other reasons because of their decisive normative status, but such a normative status is assigned by the activity of reasoning, not by the fact that they belong to an eminent domain of special objects" (Bagnoli, 2016, p. 1243).

2.3 THE VALUE OF VALUES

The term "value" is commonly used in a broad sense to encompass a wide range of moral and nonmoral ideas. For instance, in economics, value is an umbrella term that includes concepts like utility (satisfaction derived from consuming a good), efficiency (maximizing outputs with limited resources), and profitability (financial gain from economic activities). In moral philosophy, and philosophy in general, value is often synonymous with "good". A theory of value, for example, addresses questions such as "What is goodness?" and "What things are good?" This usage creates a Manichean dichotomy of prima facie distinctive normative domains: the good, which concerns what we should consider choice-worthy, and the right, which pertains to what we may or must do (Scanlon, 1998, pp. 78–79). Moral values manifest what is considered good, while moral reasons explain why a particular action is deemed right or wrong. This appears to create a contrast between intuition and justification, a tendency that dates back to at least British intuitionism, initiated by G.E. Moore with his "open-question argument" (Moore, 1903),[310] and the responses to it by philosophers such as W.D. Ross and H.W.B. Joseph.

Ross argues that rightness and goodness are entirely distinct properties: goodness is a property of motives and outcomes, while rightness is a property of acts (W. D. Ross, 1931, pp. 8–11). Acts can be right or wrong but not intrinsically good or bad; they can be instrumentally good (conducive to good outcomes), but Ross denies that instrumental value constitutes real value. Thus, according to Ross, acts themselves have no intrinsic value, while motives "can be intrinsically good or bad but never right or wrong" (W. D. Ross, 1931, p. 11). Joseph finds Ross's position absurd, arguing that rightness should not be disconnected from value and must, in some way, depend on goodness (Joseph, 1933, p. 61). Joseph also distinguishes between two senses of "right": one meaning "obligatory" and the other as a form of "goodness" tied to moral actions. To say that

[310] Moore argued that for any attempt to define "good" in terms of other properties, it remains a meaningful question to ask whether that property is truly good. This suggests that the concept of "good" cannot be fully captured by analysis or reduction: values cannot be entirely explained in terms of (natural) facts, as any attempt to do so leaves the essential question of goodness unanswered (Moore, 2004, p. 44). Moore's argument is aimed at all attempts to reduce value to other properties, whether metaphysical or natural. He specifically contended that such reductions involve a *naturalistic fallacy*, although some critics, notably Frankena (1939), have argued against this claim. Some constructivist proposals (Korsgaard, 1996b) can be seen precisely as responding to the concerns raised by the open question argument, emphasizing that moral value is best understood through the lens of reasoned agency rather than as a natural or empirically observable property.

an act is obligatory can mean either that the act is required of me personally, or that rightness is a form of goodness, which the action serves to realize, and it is the recognition of this goodness "which moves us when we do an action from a sense of obligation" (Joseph, 1933, p. 104). [311]

These themes were recently reinvigorated by T.M Scanlon, who proposes a reductive account of value in terms of reason, known as the "buck-passing" account (Scanlon 1998).[312] Scanlon suggests that something's value doesn't consist in any intrinsic property; instead, to call something "valuable" means that there are reasons to respond to it in certain ways (Scanlon, 1998, p. 97).[313] For example, if we say friendship is valuable, it's not because it has a special property of value in itself but because there are good reasons for us to appreciate and promote friendship.[314] In this view, value "passes the buck" to the reasons we have for our responses rather than being an inherent property: "to value something is to take oneself to have reasons for holding certain positive attitudes toward it and for acting in certain ways in regard to it." (Scanlon, 1998, p. 95).

---

[311] In later work, Ross refines his account by differentiating between the attributive and predicative uses of "good". The former connotate a specific function (for example, "a good knife" is good in a particular context for a specific function; it's not simply "good" in a general sense, but good as a knife) while the latter denote a proper object of satisfaction or approval (for example, when we say a good person," we mean that the person has moral virtues, not that they're good at some specific task; it means "good" in a broad moral sense). (W. D. Ross, 1939, p. 283). Still, Ross resists the idea that goodness is a relation (being worthy of approval), as this would reduce goodness to a dependent property, while he still conceive this second meaning of good as "an intrinsic property of objects" (W. D. Ross, 1939, pp. 278–279). Ewing builds on Ross's ideas, argued that goodness can be understood as something that ought to elicit favorable pro-attitudes (like approval, admiration, or desire).

[312] "At the point we have reached (the late 1940s) the intuitionists' broadly cognitivist approach to ethics was eclipsed by the onrush of non-cognitivism. Fifty years passed before these sorts of issues could again be debated with any sense of seriousness." (Dancy, 2005, p. 35). Dancy also aptly schematize this debate up to Scanlon in the following terms: Early Moore (teleological view): g → v = r (reasons defined in terms of value) 2. Later Moore: g → v → r (reasons grounded in but not defined in terms of value; value adds to the reasons given by the ground) 3. Ross: g1 → v, g2 → r (the ground for reasons differs from the ground for value) 4. Buck-passing View: g → r ↑v (being of value is having features that ground reasons).

[313] "The alternative, which I believe to be correct, is to hold that being good, or valuable, is not a property that itself provides a reason to respond to a thing in a certain way. Rather, to be good or valuable is to have other properties that constitute such reasons. Since the claim that some property constitutes a reason is a normative claim, this account also takes goodness and value to be non-natural properties, namely the purely formal, high-order properties of having some lower-order properties that provide reasons of the relevant kind. It differs from the first alternative simply in holding that it is not goodness or value itself that provides reasons but rather other properties that do so. For this reason I call it a buck-passing account (Scanlon, 1998, p. 97).

[314] See especially Scanlon's distinction between friendship and fanship, which is intended to illustrate that what we consider valuable depends on the quality of the reasons we see as "good reasons." According to Scanlon, fanship doesn't provide sufficiently strong reasons to shape one's life around it, whereas friendship offers compelling reasons that are robust enough to significantly influence and shape our commitments and actions. (Scanlon, 1998, p. 89). See paragraph 2.4 to see the different approach to the value of friendship in a constructivist proposal.

Thus, being valuable is not an independent property that gives us reasons; instead, the reasons we take ourselves to have for holding positive attitudes toward something are what make it valuable. Shared enjoyment, trust, emotional support, and mutual growth that friendships offer, give us compelling reasons to prioritize and value these relationships in our lives.[315]

Constructivism, as we already mention in the introduction, side with Scanlon in thinking that our most fundamental notion of value is broader than "the good" as this is often understood in philosophical discussion. They do so, not only by pointing out that an overemphasis on the good has had a "distorting effect on our thinking about value in general, and in particular on our view of the relation between rightness (what we owe to each other) and other values" (Scanlon, 1998, p. 79), but also by narrowing the gap between the good and the right, specifying them "in light of each other" (Richardson, 2018, pp. 33–39). Instead of using the concepts of values inside the strict dichotomy between the good and the right, constructivists understand values as a more encompassing concept that can refer to right actions and institutions (just), good objects or state of affairs (worthy of choice), and even virtue, as valuable character traits and dispositions. (Bagnoli, 2018b, p. 148; Korsgaard, 1983, p. 169).[316]

To understand how constructivists use the concept of values and its relation to the concept of right, we will proceed in three steps: first, we will explore where constructivist theories align in the fact-value distinction and within the naturalism/non-naturalism debate (2.3.1); next, we will explore more closely the distinction between intrinsic and extrinsic values to assess whether humanity itself is treated as intrinsically valuable from a constructivist perspective (2.3.2); finally, by returning on the distinction between instrumental and final end we will see how disagreements about final goods lead constructivists to adopt a form of value pluralism (2.3.3).

[315] Scanlon does not suggest that friendship—or any other type of valuable thing—should be given priority over others. Scanlon's buck-passing account doesn't rank reasons hierarchically—it only says that something is valuable if there are reasons to respond to it in certain ways. Friendship is used here merely as an example of something that can be valuable because there are reasons to respond to it in certain ways, but he isn't asserting it is necessarily a top priority compared to other valuable things.

[316] "I take it that there are three primary categories of value with which the moral philosopher is concerned: namely, the rightness or justice of actions, policies, and institutions; the goodness of objects, purposes, lives, etc.; and the moral worth or moral goodness of characters, dispositions, or actions." (Korsgaard, 1983, p. 169).

2.3.1 The fact-value distinction

Even before G.E. Moore's famous "open question argument", which challenged the reducibility of "good" to any natural property, the oddity of the "good" as contrasted with other facts was a crucial philosophical theme. David Hume's idea that one cannot logically derive an "ought" from an "is" became known as Hume's Law and was heavily used by many philosophers who argued that factual statements about the world don't necessarily imply any moral or ethical conclusions. This is the so-called *Fact-Value Distinction*, which centers around the gap between descriptive statements (what is x) and evaluative statements (if x is good or bad).[317]

In contemporary discussions, the fact–value distinction has been revitalized within the broader debate between naturalists and non-naturalists regarding the nature of values. The key question is whether values can be understood as part of the natural world—subject to empirical investigation—or whether they exist independently, beyond the reach of scientific explanation. Naturalists, such as Richard Boyd, David Brink and Peter Railton argue that moral values can be fully integrated into a scientific and naturalistic ontology: values are not mysterious or metaphysically distinct from other properties in the world, rather, they are grounded in some objective facts, like interests about human well-being (Boyd, 1980, p. 615; Brink, 1989, p. 135; Railton, 1986, p. 176).[318] By contrast, non-naturalists such as Irish Murdoch, John McDowell and Derek Parfit reject this reductionist approach, maintaining that values are irreducible to natural facts and that moral truths possess a normative authority that are *sui generis* (of their own kind), and cannot be analyzed solely in terms of natural properties or captured through empirical investigation alone (McDowell, 1996, p. 48; Murdoch, 1970, p. 86; Parfit, 2011, p. 64).[319] I will

[317] This gap, however, is different—though interconnected—with the gap between descriptive statements (about what is) and normative statements (about what ought to be). For example, if I say, "You ought to be honest" I am making a normative statement that prescribes action, but if I say "Honesty is a virtue" I am making a descriptive or an evaluative statement according to the different position on the fact-value debate. The core question here is whether the concepts of honesty and virtue can be "described" in the same way as other natural facts. This is a problem that do not pertains only to morality, but to a sheer number of evaluative statements, like "Filippo is curious".

[318] "X is *non-morally good for A* if and only if X would satisfy an objective interest of A." (Railton, 1986, p. 176).

[319] Most non-naturalists holds that moral truth are perceived through a kind of moral intuition or rational insight, like in Moore (2004). Iris Murdoch, however, challenged the traditional fact–value dichotomy by arguing that moral perception and moral understanding are themselves cognitive activities — that to see the world *justly* or *lovingly*

briefly analyze two authors from each of the opposing positions before showing why constructivists consider this debate to be misplaced.

In his 1986 work "Moral Realism", Railton starts from the assumption that since morality does not possess methodological autonomy from other sciences, the definition of "good" cannot be established a priori—otherwise, it would fall into the naturalistic fallacy; instead, it must be both revisable and reformative, meaning it should be defined in non-moral terms through an empirical-psychological reduction and validated a posteriori by its explanatory success. This is why he reduces value claims to what he calls "objective interests", attitudes of idealized versions of ourselves (Railton, 1986, p. 176). For Railton, therefore, a value claim like "X is good" means that X would fulfill the objective interests of an agent under idealized conditions of full information and perfect instrumental reason. According to Railton, objective interests provide a basis for values, rooted in empirically accessible facts, such as those about human psychology and social dynamics. This view allows moral truths to be understood similarly to scientific truths, rooted in the natural world and amenable to rational evaluation.[320]

This naturalistic reductionist account explains what it means for something to be good for someone, but it still leaves open the question of how this carries normative force. To address

---

is already a kind of moral knowledge, not a subjective response added to otherwise neutral facts (Murdoch, 1970, p. 86). Building on Murdoch's insight, John McDowell, in *Mind and World* (1996), proposes that values are part of our conceptual framework and are perceived through experience, much like secondary qualities (e.g., color). He argues that values are not merely projections of subjective attitudes but are part of the "space of reasons" we inhabit as rational beings. For McDowell, understanding values involves an interpretative process that is not reducible to empirical observation but is also not entirely detached from reality. This approach challenges both strict naturalism and moral intuitionism, proposing that values are objectively present within our rational understanding of the world, though not as straightforwardly observable as physical facts (McDowell, 1996, p. 48).

[320] There is an important distinction to be made between Railton's naturalism and Aristotelian naturalism, a tradition that can be traced back at least to Elizabeth Anscombe's paper *Modern Moral Philosophy* (1958), where she criticized the dominance of consequentialist and deontological ethics and called for a return to virtue ethics inspired by Aristotle. A more refined and systematic development of this approach is found in Philippa Foot's *Natural Goodness* (2003), where she argues that human beings, like other living organisms, have a "form of life" with natural functions that contribute to their flourishing. Michael Thompson, in *Life and Action* (2008), further expanded on this idea, emphasizing the biological and teleological aspects of moral evaluation by grounding ethical norms in species-specific life forms. While both Railton's naturalism and the Neo-Aristotelian naturalism reject non-naturalist views that treat moral properties as wholly separate from the natural world, they do so in very different ways. Railton's reductionism seeks to ground moral truths in objective facts about human well-being, arguing that moral claims can be analyzed in terms of what would be rationally approved under conditions of full information, whereas Aristotelian naturalism remains normatively rich, preserving the idea that moral norms are rooted in the essential nature of human life rather than in contingent psychological states or social agreements.

normativity, Railton examines our ordinary use of "ought" and offers an example involving roof planks (Railton, 1986, p. 185). Suppose we construct a roof using planks that are too small to support the expected weight. When the first heavy snowfall arrives and causes the roof to collapse, we naturally say, "We ought to have used larger planks." Railton interprets this normative statement as reducing to a conditional claim: "If we want our roof to remain stable, we must use larger planks." In this view, the motivational force of normativity arises from instrumental reasoning and preexisting value commitments. Normative statements, then, express rational imperatives grounded in an agent's goals and the factual conditions required to fulfill them, but they should not be understood in universalistic terms; otherwise, normativity would be irreducible to the natural entities studied by science, thus reinstating the fact-value dichotomy. However, knowing what is good for a particular individual is not enough; we also seek to understand what is good for everyone or what is good "all things considered". To address the latter issue, Railton introduces the "social point of view"—a perspective that takes into account the interests of all individuals. From this standpoint, moral obligations are determined by what "would be rationally approved of were the interests of all potentially affected individuals counted equally under circumstances of full and vivid information" (Railton, 1986, p. 190). While Railton acknowledges that this approach entails a form of consequentialism at the normative level, he claims that it differs from traditional utilitarianism precisely due to his underlying account of value.[321]

[321] Classical utilitarianism, developed by Jeremy Bentham, located moral value in the maximization of pleasure and the minimization of pain: the only intrinsic good is individual utility, understood as the quantitative balance of pleasure over suffering (Bentham, 1789). Only by starting from this foundation it was possible to move to the social dimension, since the right action is that which produces "the greatest utility of the greatest number." John Stuart Mill refined this framework by introducing a qualitative distinction among hedonistic pleasures, yet he retained its individualistic core: moral rightness is determined by how much *happiness* an action promotes (Mill, 1861). Therefore, while for classical utilitarianists individual utility is ultimately directed toward collective utility, this collective dimension is grounded in a fundamentally individualistic epistemology and a subjectivist determination of value. In the late nineteenth and twentieth centuries, Henry Sidgwick and R.M. Hare further systematized utilitarian ethics, introducing a more systematic foundation grounded in rational principles rather than mere subjective preference, though in sharply different directions (Hare, 1992; Sidgwick, 1874). Sidgwick, in *The Methods of Ethics* (1874), developed a comprehensive and systematic account of utilitarianism grounded in moral realism and rational intuitionism. He introduced intuitionism because he believed that neither Bentham nor Mill had satisfactorily explained how the moral requirement to sacrifice one's own interest for the sake of the general good could have genuine normative force, especially given their psychological assumption that individuals always act in pursuit of their own self-interest. For Sidgwick, such a moral demand could only be justified if rooted in self-evident ethical principles accessible through reason—most notably, the principle of universal benevolence, which serves as the rational foundation for utilitarian ethics. Hare, by contrast, offered a fundamentally anti-realist and non-cognitivist reinterpretation of utilitarianism through his theory of "universal prescriptivism" (Hare, 1992). He insisted that moral judgments do not describe objective truths but function as prescriptive commitments—a significant departure from Sidgwick's rationalist realism. Hare also distinguished between two levels of moral reasoning: the "intuitive level", where moral agents rely on established rules and heuristics to guide everyday decisions, and the "critical level", where

Railton's proposal ultimately links morality—conceived as the progressive social maximization of individual interests—to rational feedback and social control dynamics, where the rational element is center around and ideal of a fully informed and rational agent, whose standpoint determines what genuinely serves an individual's well-being. In this sense, Railton proposal is a form of non-reductivism, since moral facts are naturalistically grounded yet not reducible to mere empirical regularities: they depend on what such an ideal agent would endorse under conditions of complete information and rational coherence. In this view, the rigid fact–value distinction is ultimately misguided because moral facts (or values) are continuous with empirical facts. They can be studied, explained, and justified using the same scientific and rational methods we apply to other areas of inquiry. However, this approach appears ill-equipped to determine the permissibility of certain interests before social feedback occurs. This issue is especially evident in the problem of maximization: why should rational agents prefer the implementation of an egalitarian principle that reduces inequalities? It is not contradictory to imagine fully rational agents choosing not to redistribute their wealth to lower-income groups in order to reduce inequality. Railton's only apparent response to this issue is that even the most privileged individuals might come to recognize that extreme socio-economic inequality can, in the long run, lead to problems that affect them, such as uprisings and social unrest. This exposes a theoretical confusion about the nature of normativity, which, according to many anti-naturalists, is not properly accounted for in Railton's framework. However, this is not a mere shortcoming of Railton's arguments, but rather a

---

those rules are evaluated or overridden through abstract, universalizable reasoning (Hare, 1992, pp. 25–26). Railton draws from both, but his peculiarity lies in reintroducing moral realism into utilitarian ethics through a naturalistic framework. He does so, as we saw, by grounding moral value in objective facts about what would promote the interests and well-being of fully informed and rational agents (Railton, 1986, p. 190). In doing so, Railton reformulates Hare's two-level structure to address one of the classical challenges to utilitarianism—the "alienation objection" (Stocker, 1976)— arguing that a morally committed utilitarian should not constantly deliberate in explicitly utilitarian terms (Railton, 1984, p. 137); instead, Railton distinguishes between two dimensions of moral reasoning:1) Objective consequentialism – the standard by which actions are truly right or wrong, determined by their actual consequences and impact on overall well-being; 2) Subjective motivation – how individuals think, feel, and act in practice, guided by personal attachments and ordinary motivations. Accordingly, the "sophisticated consequentialist", in Railton's view, is someone who 1) is committed to living a consequentialist life overall, aiming to promote the best outcomes in the long run; 2) yet does not continually engage in explicit consequentialist reasoning in daily life. "[W]hile [the sophisticated consequentialist] ordinarily does not do what he does simply for the sake of doing what's right, he would seek to lead a different sort of life if he did not think his were morally defensible." (Railton, 1984, p. 151). In short, Railton holds that utility, defined as human well-being, is the sole intrinsic value—namely, what would be desirable for a fully informed and rational agent; however, he rejects the classical utilitarian view that this must be framed in strictly subjectivist or hedonistic terms, opting instead for a more objective, naturalistic and monist conception based on an idealized agent.

methodological impossibility: no naturalist approach, insofar as it adopts reductivist empirical methods, can fully capture normativity

This type of rejection is precisely one of the main arguments that Derek Parfit's developed in *On What Matters*, a work spanning three standalone volumes;[322] according to Parfit these naturalist theories fail to account for the phenomenon of values and normativity, which, in his view, is what truly "matters." Parfit starts his critiques with naturalist views of reason for being necessarily internalist, a position he considers, as mentioned in the previous section, self-defeating: if internalism were true, reasons could not guide us independently of the desires we already possess. This would reduce normative claims to mere analytic or tautological statements, stripping them of any genuine normative force. In contrast, substantive normative reasons, according to Parfit, cannot be reduced to naturalistic notions but must be understood as primitive and unanalyzable. They cannot be explained in terms of causal or empirical facts since the only valid analytic understanding of normativity is one that treats it as entirely irreducible, akin to other a priori truths. Just as the mathematical statement $3 \times 4 = 12$ holds independently of any natural facts we might use to explain our reasoning, moral truths—such as the distinction between just and unjust actions—exert normative force on us regardless of any empirical considerations. This is what Parfit calls the "Normativity Objection" (Parfit, 2017, p. 72).[323]

Parfit identifies two positions that naturalists can take in response to this issue. Hard Naturalists argue that since all facts are natural, irreducibly normative claims are unnecessary: all normative discourse can be eliminated or reduced to purely descriptive terms. Soft Naturalists, on the other hand, acknowledge the need for normative claims to guide decision-making and action but maintain that true normative statements ultimately describe natural facts. While Parfit implicitly challenges Hard Naturalism by suggesting that normative concepts and statements are essential and cannot simply be eliminated or reduced to naturalistic descriptions, his primary focus is on demonstrating why Soft Naturalism is also flawed. He argues that if there are no irreducibly

[322] I will primarily focus on Volume III, as it was published a decade after the first two volumes and incorporates Parfit's engagement with updated versions of naturalist theories, including Railton's revised position.

[323] "I have just stated the simplest and most straightforward objection to Normative Naturalism. According to this Normativity Objection: Irreducibly normative, reason implying claims could not, if they were true, state normative facts that were also natural facts." (Parfit, 2017, p. 72).

normative truths, then normative beliefs lose their genuine normative force—they would not actually help us make better decisions or act well. This argument, which he calls the "Triviality Objection", exposes a key weakness in Soft Naturalism: without irreducibly normative truths, normativity itself becomes empty, failing to fulfill its supposed practical function. If natural facts could express normative truths, this truths would be normatively trivial, meaning they would provide no substantial normative guidance or insight (Parfit, 2017, p. 8).[324]

Parfit attacked explicitly Railton and his attempt to ground moral truths in natural facts about human well-being, precisely because these facts cannot be understood in terms of irreducible normative facts. Railton's initial response to Parfit's objections attempted to address the issue by appealing to the complex roles or functional descriptions that certain natural properties might fulfill in our normative discourse. In other words, he suggested that even if normative facts are ultimately natural facts, they play a special role in our reasoning, guiding action and moral deliberation. This solution is intended to show that normative statements are not trivially descriptive, but meaningfully prescriptive because of the functions they serve (Parfit, 2017, p. 95).[325] However, Parfit was not convinced by this defense and argued that merely assigning a special role to natural facts does not give them genuine normative force. That is, just because certain natural properties (e.g., well-being, social cooperation) function as standards in moral reasoning does not mean they actually possess irreducible normative authority—they remain just descriptive facts about the world (Parfit, 2017, p. 97). Railton later responded to these objections by expanding his view of normative naturalism, arguing that one can justifiably believe in the existence of certain normative truths, as this belief would not introduce any mysterious elements

[324] "Soft Naturalism, I argued, could not be true. If there were no irreducibly normative truths, our normative beliefs could not help us to make good decisions and to act well. I called this argument the Triviality Objection, but this name is in one way misleading, since it would not be trivial if there were no irreducibly normative truths" (Parfit, 2017, p. 8). Therefore, even in this modest formulation, Soft Naturalism is not a form of non-reductionism: it acknowledges the usefulness of normative language, yet continues to treat normative truths as ultimately reducible to natural facts. Non-reductive naturalism, by contrast, goes further: it holds that normative facts are natural but irreducible—that is, they possess an autonomous dimension that cannot be fully explained in purely scientific or naturalistic terms.

[325] "In the second half of his paper, Railton suggests some answers to this objection to Soft Naturalism. Naturalists, he writes, can explicate certain concepts in terms of a complex role or 'job description' to be satisfied by whatever would satisfy this concept. This remark is like my claim that some concepts have a gap that is wait ing to be filled. Molecular kinetic energy, we could say, satisfies the role or job description of being the property that can have certain effects, by causing us to have certain sensations, melting solids, turning liquids into gases, etc. We could similarly claim that some natural property, such as the property of minimizing suffering, might satisfy the job description of being the property that makes acts right." (Parfit, 2017, p. 95).

into our ontology: naturalists can therefore accept non-natural properties in a nominal or linguistic sense and tolerate the use of irreducible normative concepts to reflect on what matters and to deliberate about what to think or do (Parfit, 2017, p. 100).[326]

Parfit recognized that Railton's expanded view effectively addressed his objections to normative naturalism, leading them to a significant point of agreement. [327] He highlighted this friendly passage from Railton as evidence of their shared understanding: "Derek Parfit's response to my paper is, to me, immensely heartening. We are indeed climbing the same mountain. Moreover, despite our different starting points, our paths have brought us increasingly close. Of course, the geometry of mountain peaks in itself tends to produce this result— so long as all parties are actually making progress toward the summit. And that is what is so heartening: Parfit and I each can now recognize the other as having worked his way upward, away from mutually uncomprehending beginnings—he a Non-Naturalist who thought Naturalists like myself were simply unable to see the nature of the normative, and I a Naturalist who thought the idea of 'non-natural properties' smacked of Plato's heaven—and worked our way toward a common idea of how to understand the place of the normative in the natural world. In this brief reply on my part I hope to say something about the nature of that convergence, and also about some challenges that lie between us and the summit" (Parfit, 2017, p. 113).

In the end, the supposed divide between naturalism and non-naturalism appears to be more of a misunderstanding than a genuine opposition if even the two most prominent figures in the debate, Railton and Parfit, ultimately recognized that they were climbing the "same mountain" rather than standing on opposite sides. Their eventual convergence suggests that the dispute may have been exaggerated or mischaracterized, and that the core issue lies more in the framing of normativity within a broader philosophical and metaethical landscape that goes beyond a strict

[326] "Commenting on these parts of my view, Railton writes: Soft Naturalists . . . can accept non-natural properties in a nominal or linguistic—and to that extent non-ontological— sense. Soft Naturalists also allow us to talk meaningfully and truthfully in terms of normative concepts—making reasons claims in which normative predicates figure. True reasons claims can be called normative facts by the Soft Naturalist in one familiar sense of this term—a fact is the content of a true statement or proposition." (Parfit, 2017, p. 100).

[327] "I cannot find points of agreement between views which assert and views which deny that we have reasons to care about some things. But when I wrote Part Six of On What Matters, I misunderstood some of the people whose meta-ethical views I rejected. Two such people are Railton and Gibbard. I now believe that, as Railton and Gibbard have separately suggested and I shall later try to show, the three of us have resolved our main meta-ethical disagreements. We hope that others will reach similar conclusions" (Parfit, 2017, p. 2).

distinction between facts and values. This perspective is not new, as other scholars had already pointed out the potential for reconciliation long before Parfit's Volume III explicitly sealed the debate. One notable example comes from Hilary Putnam, who, in his 1982 article, titled precisely *Beyond the Fact-Value Dichotomy*, challenged this rigid distinction. Putnam argued that even in science, theoretical choices involve evaluative judgments—scientists do not merely record objective facts but must decide between competing theories based on criteria such as simplicity, coherence, and explanatory power. He illustrated this by discussing Einstein's theory of gravity, which was preferred over alternatives (e.g. Nordström's scalar theory) not because of some purely factual superiority, but because it provided a more elegant and unified account of physical laws (Putnam, 1982, pp. 6–7).[328]

Constructivists also view the fact-value distinction as largely a relic of outdated metaphysical commitments. While values are part of the ontology of our world, they are not inherent features waiting to be discovered; rather, they are constructed through the rational deliberation and procedural standards of agents (Korsgaard, 1996b, p. 112).[329] This means that values are not objective in the sense of existing independently of human thought, but attain objectivity within a shared rational framework (Rawls, 1997). For instance, principles of justice may take specific human needs and social contexts into account, yet they are not dictated solely by these empirical facts. Instead, they emerge from deliberative processes concerning how rational beings ought to treat one another. There is no "moral queerness" here—values do not stem from

[328] "When theory conflicts with what is taken to be fact, we sometimes give up the theory and sometimes give up the "fact"; when theory conflicts with theory, the decision cannot be always made on the basis of the known observational facts (Einstein's theory of gravitation was accepted and Whitehead's alternative theory was rejected fifty years before anyone thought of an experiment to decide between the two). Sometimes the decision must be based on such desiderata as simplicity (Einstein's theory seemed a "simpler" way to move from Special Relativity to an account of gravitation than Whitehead's), sometimes on conservativism (momentum was redefined by Einstein so that the Law of the Conservation of Momentum could be conserved in elastic collisions) ; and "simplicity" and "conservativism" themselves are words for complex phenomena which vary from situation to situation. When apparent observational data conflict with the demands of theory, or when simplicity and conservativism tug in opposite directions, trade-offs must be made, and there is no formal rule or method for making such tradeoffs. The decisions we make are, "where rational, pragmatic", as Quine put it." (Putnam, 1982, pp. 6–7).

[329] "Values are not discovered by intuition to be 'out there' in the world. Good maxims are intrinsically normative entities, but they are also the products of our own legislative wills. In that sense, values are created by human beings. Of course we discover that the maxim is fit to be a law; but the maxim isn't a law until we will it, and in that sense create the resulting value. This is what Pufendorf means when he says that moral entities are produced by imposition, and that imposition in turn is a matter of making laws.24 The form of realism I am endorsing here is procedural rather than substantive realism: values are constructed by a procedure, the procedure of making laws for ourselves" (Korsgaard, 1996b, p. 112).

an independent moral realm but from rational agency itself, and their normativity arises not from natural properties or non-natural essences but from the rational and procedural commitments of agents engaged in reasoned discourse (Forst, 2017b).

While constructivists deny that values are simply reducible to facts, they still acknowledge that values are not entirely disconnected from factual considerations. Human vulnerability, social cooperation, and interdependence, for instance, are factual aspects of human life that inform moral reasoning (O'Neill, 1988). However, constructivists argue that values are not determined by these facts alone but are instead constructed through rational deliberation on how to respond to them. Therefore, while facts about human nature or society may influence values, they do not dictate them; instead, values arise from rational agents actively engaging with these conditions (Bagnoli, 2016, p. 1235).[330] This is the practical standpoint from which values are constructed, not in isolation, but through interpersonal reasoning and engagement with others. This intersubjective basis for moral values contrasts with the idea of moral truths as fixed, external realities, offering a more flexible yet objective grounding in shared rational processes (Habermas, 1998).

In short, constructivists reject a rigid fact/value dichotomy that implies that values exist independently of human evaluative frameworks. Instead, they argue that values are constructed through rational intersubjective processes and reflect the principles that rational agents would endorse under specific conditions. This perspective challenges the idea that moral and normative truths are external and waiting to be discovered, suggesting instead that they emerge through human reasoning and shared deliberative practices. By emphasizing the role of rational deliberation constructivism highlights the fluid and dynamic nature of values, which can evolve over time as societies progress and as individuals reflect on their commitments, without collapsing in form of moral relativism: values are not arbitrary but are shaped (and not imposed) by our constitutive principles of reasoning.

[330] "Constructivism does not deny that there are facts or objects that can be called 'moral' or, more broadly, "normative". What constructivism denies is that such facts and objects are "moral" or "normative" in themselves, prior to and independently of the intervention of reason. This denial preserves the autonomy of reason and it is also meant to explain why reason is effective, that is, practical. The intervention of reason is necessary in order to transform knowledge of some objects into "practical knowledge", namely, knowledge of normative propositions that are binding in the first person and thus carries importance for us, practical subjects. The intervention of reasoning consists in "making up normative propositions" which is exactly the process that goes under the name of construction." (Bagnoli, 2016, p. 1235).

2.3.2 Moral values: Intrinsic or extrinsic?

Metaethical questions about values do not stop simply by determining if value is or is not mind-dependent (ontological questions): even if we accept that value is mind-dependent, it is still reasonable to ask where the source of something's value lies (axiological questions). Specifically, it is genuinely possible to ask whether a thing's value is *intrinsic* (valuable in itself) or *extrinsic* (valuable because of its relationships with other things). [331] Even if considering a value both mind-dependent and intrinsic may seem counterintuitive at first because we often associate "intrinsic value" with objective, mind-independent worth, the two are logically independent. Consider a classic example of value, which is happiness; many people view happiness as intrinsically valuable, something worth pursuing for its own sake and not for its effects or relations to something else. However, this value is still mind-dependent because happiness itself is a mental state, dependent on our subjective experiences and perceptions. Contrasted with intrinsic values there is extrinsic value, which refers to the value something has because of its relationships or effects in relation to something else. Action Comics #1 (the volume that introduced Superman in 1938) is extrinsically valuable in virtue of having the extrinsic property of rarity. The source of its goodness, in this regard, is not contained within itself, it is located outside of the entity and while the comic in itself is of course a mind-independent object, the rarity connected to it is a mind-dependent property: after a mass-extinction of human beings, that comic, which still exist and remain in the world, has no more the property of rarity connected to it.

Dividing scholars in value theory into internalists and externalists may be a useful distinction, but within these broad categories, there is significant diversity among theorists who propose different accounts of moral value. Among value internalists, various intrinsic properties have been identified as the source of moral value; while some of these properties are attributed exclusively to humans, others extend to non-human beings as well. These might include: 1)

[331] Sensen distinguish between four possibilities regarding the nature of value, which mix ontological and axiological questions. "Value could be a distinct metaphysical property, in the sense that it is a property an object possesses even if it were to exist in isolation. Second, value could be a relation between two objects, e.g., a relation of usefulness or fittingness. Third, value could be subjective, in the sense that it is what a subject does value (e.g., pleasure, happiness). Finally, value statements could be nothing more than a (rational) prescription that commands what one should value." (Sensen, 2011, p. 29). For Sensen, Kant uses the concept of value only in this fourth sense. "If the faculty of reason deems a maxim or an action as necessary, Kant can say that this maxim or action 'has value', without ascribing a metaphysical property to the action. 'Has value' is merely another way of saying 'should value' or 'would value if he were fully rational'." (Sensen, 2011, p. 30).

rationality and practical reasoning (Gewirth, 1980); 2) the capacity to value something (Rolston III, 1994); 3) having an integrated nature with agency (Schlosberg, 1999); 4) sentience experience and having interests (Singer, 2009); 5) the possibility of experiencing frustration (Nussbaum, 2013). In contrast, value externalists underline that moral value arises from certain relationships. Some of the proposed relational sources of moral value include: 1) social bonds (MacIntyre, 1981); 2) ecological holism and biodiversity (Callicott, 1989); 3) the possibility of Recognition (Honneth, 1995); 4) the constitutive relationship (Pelenc et al., 2016); 5) the complex interconnectedness with everything (Van Jaarsveld, 2020); 6) the relationship of ancestrality (Kramm, 2020).

Kantians are usually categorized inside the current of intrinsic properties, especially for the idea of the intrinsic value of humanity often associated with them. For example, Nandi Theunissen, influenced by David Velleman and Katja Maria Vogt's work, develops a relational proposal that directly challenges the Kantian view of humanity's absolute and unique value: "Since value is a relation, it cannot be as Kant proposes" (Theunissen, 2020a, p. 3).[332] Contrary to the idea that human value is incomparable to anything else, she argues that our value is continuous with that of other valuable things and presents lines of thought suggesting that value (or goodness) is affecting: it helps, serves, and enriches. Starting from the Socratic notion that the good is inherently beneficial—expressed in contemporary terms as "good for someone"—Theunissen posits that if people are bearers of value, our value must also follow this principle and should therefore be explained through reciprocal relations or interdependence. For instance, as daughters, teachers, or friends, we benefit others by being integral parts of relationships with them.

However, ultimately Theunissen expand this relational account by opening the possibility that we can stand in a valuable relationship with ourselves: a person matters because she values herself in a unique way—one that enables her to lead a flourishing life (Theunissen, 2020a, p. 97).[333] Yet, this form of self-valuation does not justify the Kantian claim that human beings possess

[332] "I do not find it plausible to suppose that Kant was a normative realist about human value, as if our value were a mind-independent property of the kind that is theorized by G.E Moore and those following him. This fits ill with Kant's broader philosophical commitments. In calling human beings absolutely valuable, I take Kant to mean that human beings are such that they categorically must be valued: in other words, valued independently of their capacity to stand in relations of benefit to something or someone" (Theunissen, 2020a, p. 3). The constructivist reading of Kant challenges this interpretation, rejecting the idea that Kantian value of humanity relies on an intrinsic, mind-independent property (Sensen, 2011, pp. 133–134).

[333] "If valuing is the basis of our value, and bearing the relevant relationship to ourselves is the explanation of our value, then the proposal is that we are of value because, as valuers, we are able to be good for ourselves in the

intrinsic value in an independent, non-relational sense, and this is why Theunissen question whether anything can be valuable in the way Kantians claim. As she puts it: "My dissatisfaction with this framework arises from a suspicion that anything could be valuable in the way Kantians take human beings to be of value. That is, human beings cannot be of value independently of their propensity to stand in relations of benefit to human beings or other beings because these relations are the very essence of value" (Theunissen, 2020a, p. 135).

This critique also extends to Kantian constructivists and is arguably one of the most common objections they face (Shafer-Landau, 2003, p. 42), namely that their framework implicitly relies on the value of humanity to ground moral obligation, thereby "committing to moral realism" (Bagnoli, 2016, p. 1229) and assuming a form of intrinsic value. At the same time, however, some constructivists, most notably Korsgaard (1996) were criticized for locating the normative source of moral obligations in the agent's self, rather than the value of humanity (G. A. Cohen, 1996; Skorupski, 1998; LeBar, 2001; J. Gert, 2002; Skidmore, 2002; Wallace, 2009). This twofold critique can be addressed by more clearly articulating the constructivist position with respect to both horns of the dilemma. Korsgaard agrees that humanity holds value (Korsgaard, 1996b, p. 121)[334], but her point is that the value of humanity can be explained by understanding what it means to value something, rather than by assuming that humanity's value is a starting point for moral reasoning. As reflective beings, we practically rely on reasons for acting, and these reasons for action are rooted in particular practical identities—that is, in conceptions of ourselves that we hold as valuable (Korsgaard, 1996b, p. 101);[335] nonetheless, any particular practical identity is ultimately grounded in our shared humanity (Korsgaard, 1996b, p. 129). Therefore, to value any specific practical identity, one must first value one's humanity as its source: to act at all on the

sense that we are able to lead good lives – lives that are good for us. More fully spelled out, people are of value in virtue of a capacity to value (intrinsic property), where that capacity, when exercised, makes people good for something (relational property), as on the view I favor, it makes people able to lead good lives" (Theunissen, 2020a, p. 97). Kantians constructivists and Theunissen agrees on the fact that without valuers, there are no values, and they also agree on the reflexive component that explicit the specific type of valuable relationship with ourselves. The difference is that, as we will see in the next pages,

[334] "Valuing humanity in your own person rationally requires valuing it in others" (Korsgaard, 1996b, p. 121).

[335] "To treat our human identity as normative, as the source of reasons and obligations, is to have… [a] moral identity.". (Korsgaard, 1996b, p. 101).

basis of reason, rational agents must inherently value their own humanity (Korsgaard, 1996b, p. 123).

While Korsgaard's argument may justifiably be critiqued for not logically following as she expects (Sensen, 2011, p. 64),[336] constructivists have further refined the position, clarifying that valuing our humanity, autonomy, or capacity for good will is a constitutive element of the act of valuing itself. We do not respect others merely because they share the same capacities as ourselves, but because one must "be constrained to follow the law springing from one's own reason" (Sensen, 2011, p. 134). Therefore, the critique of constructivism as a form of intrinsic realism regarding human value is misplaced. This critique assumes that the moral domain consists of special, pre-existing objects of value. In contrast, for constructivism, there is no moral domain of objects independent of reasoning. The very metaphor of 'constructions' is introduced to affirm the idea that reasoning precedes its objects (Korsgaard, 1996b, pp. 38–40).

Although constructivists do not claim that values are inherent in the world, they believe that values can achieve a kind of objective validity through universally applicable rational processes. Kantian constructivists might say that the value of respecting individuals as ends in themselves is not intrinsic to the universe but is a necessary conclusion of rational deliberation; this value gains "intrinsic" status within the constructed framework of rational moral agency (Korsgaard, 1996b, pp. 120–131; Sensen, 2011, pp. 188–191), but remains (one of) the result of rational deliberation, not the assumed starting point of it as Theunissen characterize it. In this view, intrinsic and extrinsic values are therefore intertwined, as what is valued is shaped by the reasons and justifications that arise from our rational agency. Thus, in a similar vein to Scanlon (1998, p. 79), constructivists reject the notion of value as an inherent property that exists independently of our interactions and rational deliberations. Instead, they favor an understanding of value that emphasizes its relational and context-dependent nature. Yet, this adaptability to context does not imply the acceptance of arbitrary values or a pre-determined, fixed hierarchy of values, as we will explore in the next section.

[336] In the next chapter, we will explore Sensen's critiques of Korsgaard in greater detail, including the insightful analogy that just because something is perceived as funny only by humans, it does not mean that humans are inherently funny.

2.3.3 Final ends, Value Pluralism and Incommensurability

After clarifying the distinction between intrinsic and extrinsic value within a constructivist framework, by focusing on the reasons we value things, the source of their goodness (in short, *why* we consider them valuable), it is essential to examine the distinction between ends and means, which concerns the ways in which we value things, the roles they play in our evaluations (in short, *how* we attribute value to them). This distinction allows us to further develop the discussion on ends, desires, and instrumental rationality introduced in section 2.2.3.

Ends are things we value for their own sake (Korsgaard, 1983, p. 170; Richardson, 1994, p. 50).[337] Means instead are things we value for the sake of other things, for example as causal means to other things we value for their own sake. If I value my cat for her own sake or as an end, I regard her with direct esteem as the thing she is. In the event of her loss, no other cat could serve as an appropriate replacement, because I valued her in her own right and we shared a unique bond, which could not be equivalently shared with a different cat. Instead, money is a means, because I value it only as long as it can be used to purchase goods or services that I value as end (that can have intrinsic or extrinsic values).

Korsgaard was the first to highlight the importance of distinguishing this fourfold dimensions of valuing (Korsgaard, 1983, p. 170).[338] This distinction reveals that some ends do not need to be intrinsic to be objectively valuable (Korsgaard, 1983, p. 195);[339] they can still hold

[337] "An end is an aim of an action. It is something for the sake of which an action is to be done. The normative aspect is called in the phrase 'for the sake'." (Richardson, 1994, p. 50).

[338] "A common explanation of the supposedly contrasting kind, intrinsic goodness, is to say that a thing is intrinsically good if it is valued for its own sake, that being the obvious alternative to a thing's being valued for the sake of something else. This is not, however, what the words "intrinsic value" mean. To say that something is intrinsically good is not by definition to say that it is valued for its own sake: it is to say that it has its goodness in itself. It refers, one might say, to the location or source of the goodness rather than the way we value the thing. The contrast between instrumental and intrinsic value is therefore misleading, a false contrast. The natural contrast to intrinsic goodness-the value a thing has "in itself"-is extrinsic goodness-the value a thing gets from some other source. The natural contrast to a thing that is valued instrumentally or as a means is a thing that is valued for its own sake or as an end. There are, therefore, two distinctions in goodness. One is the distinction between things valued for their own sakes and things valued for the sake of something else-between ends and means, or final and instrumental goods. The other is the distinction between things which have their value in themselves and things which derive their value from some other source: intrinsically good things versus extrinsically good things. Intrinsic and instrumental good should not be treated as correlatives, because they belong to two different distinctions." (Korsgaard, 1983, p. 170).

[339] "The distinction between a thing that is intrinsically good and a thing that is extrinsically good yet valuable as an end allows for the possibility that the things that are important to us have an objective value, yet have that value because they are important to us. Objective goodness is not a mysterious ontological attribute. The things that are

objective value even if they depend on external factors, such as individual interests. [340] Later constructivists expanded on this idea, arguing that both intrinsically and extrinsically good ends, but especially intrinsically valuable ends (sometimes referred as final ends, final goodness or final values)[341] are not merely a matter of personal preference but must be justifiable upon rational reflection (Bagnoli, 2018b, p. 149; Forst, 2024, p. 144).[342] These ends are not exempt from rational scrutiny simply because they are considered "final", as some instrumentalists might argues, since very often they are not definite enough to serve as the starting point for means-end reasoning (Kolnai, 1961, p. 201; Millgram, 2005, p. 61; Richardson, 2013, p. 235; Bagnoli, 2016, p. 1230).[343]

---

important to us can be good: good because of our desires and interests and loves and because of the physiological, psychological, economic, historical, symbolic and other conditions under which human beings live." (Korsgaard, 1983, p. 195).

[340] However, Korsgaard herself has been criticized for confusing two different concepts of goodness in her analysis. Rabinowicz and Rønnow-Rasmussen argue that her claim, "there are final values that are not intrinsic," is flawed because she mixes up two interpretations of the source of value distinction (Rabinowicz & Rønnow-Rasmussen, 2000, pp. 34–37) One interpretation focuses on the features of an object that determine its final value—its "good-making" properties; the other concerns the foundational reasons for valuing something, such as our "preferential attitudes." I believe that this reading misinterprets the constructivist framework: when Korsgaard talks about the sources of value, whether intrinsic or extrinsic, she's referring to the qualities that give us reasons to value something as good. There is no rigid distinction between the properties that an entity has in order to be considered good and my attitudes towards it that makes it worthy of choice. As she puts it "My concern here is not with what constitutes moral worth or moral goodness but with the second category-with goodness as a feature of ordinary ends and purposes, states of affairs, objects, activities, and other things-that is, with the kind of goodness that marks a thing out as worthy of choice." (Korsgaard, 1983, p. 169)

[341] "The simple notion of an end cannot bear the full burden of the evaluative status of the goal: we need to invoke the idea of a final end. […] A final end., sought for its own sake, is valued in itself. This implies something about the evaluative (and not just the causal) context. Since final ends provide fundamental reasons for action, the practical implications of a final end must go beyond ruling certain options out. A final end must in addition indicate something positive about when the agent would act." (Richardson, 1994, p. 53).

[342] "On some theories, the distinction between instrumental and final goods collapses on, or is equated to, another distinction between intrinsic and extrinsic goodness. Instrumental goods are also held to be extrinsic goods, and final goods intrinsic goods. However, the distinction between intrinsic and extrinsic goodness concerns the sources of values, while the distinction between final and instrumental goods concerns the way we attach value to items. To bear in mind this distinction allows us to appreciate that there are things valued as extrinsically good, and yet as final ends. For instance, painting or horseback riding might be final ends because of the interest we take in them. […] They are not unconditionally good, since their being good is relative to us, but they are nonetheless rationally justified on the basis of their desirability" (Bagnoli, 2018b, p. 149).

[343] "[R]easoning, is in some important sense independent of what reasoning is about, and of what reasoners happen to desire or value, but not independent of who they are." (Bagnoli, 2016, p. 1230). This is why the key-feature of reasoning is its transformative function, as we have underline in paragraph 2.2.3, and this transformation includes also the ends they endorse. "Undertaking reasoning is a transformative experience for practical subjects. Their desires, interests, and concerns are not fixed but change also in virtue of rational deliberation. The reasons practical subjects construct could may be rather different from the ones they thought they had before engaging in reasoning. Subjects may end up with reasons whose content radically differs from the particular interests that they wanted to protect when they entered reasoning (Bagnoli, 2016, p. 1243).

As we anticipated in the previous section, constructivists believe that, through deliberation, we can refine our ends, arriving at more justified or fitting ones. This deliberation process is open-ended: it includes examining and questioning our goals and purposes, allowing for possible revisions based on deeper insights or broader perspectives (Richardson, 1994, p. 69). We are not passive recipients of our final ends; rather, we actively shape, refine, and redefine them through our rational agency, which serves as the source of value. Our engagement with values is dynamic, and we exercise our autonomy to choose and pursue ends that align with our understanding of what is worthwhile.

For instance, someone might choose painting as a final end, not because painting inherently possesses some intrinsic value, but because it helps them achieve broader personal or rational goals—such as self-expression, creativity, or emotional fulfillment—that are meaningful within their life context. However, this way of proceeding might introduce a potential issue: since individuals' rational processes and personal goals can vary widely, this inevitably leads to disagreements about what constitutes value. This poses a challenge, as such disagreements raise questions about the objectivity and universality of value claims, which some critics find problematic, especially in a framework that purports to provide a rational basis for moral judgment.

There are two main responses to the problem of value disagreement: monism and pluralism. Monists argue that, despite apparent disagreements, there is ultimately a single fundamental value and a single fundamental principle that can resolve conflicts. This position is often associated with moral realism (values exist in the world and can be discovered through experience or reason) and monists believe that disagreements about values stem from factors such as lack of knowledge or misunderstandings, and that with sufficient rational deliberation a consensus can be reached (Boyd, 1988; Shafer-Landau, 1994). Utilitarianism for example claims that all moral values can be reduced to a single value, which is utility (Hare, 1992).[344] While monism offers the advantage of providing clear decision-making criteria and consistency in moral

[344] The definition of utility varies across different versions of utilitarianism, as we briefly mentioned while discussing the fact-value distinction in section 2.3.1 (especially footnote n. 321). Classical utilitarians like Jeremy Bentham and John Stuart Mill equated utility with pleasure or happiness, while contemporary utilitarians define it more broadly in terms of well-being or preference satisfaction (Harsanyi, 1977; Hare, 1992; Singer, 2011). Richardsons, which follows Rawls, believes that even the "preference-based idea of utility cannot guide people's deliberations (in its capacity as a good) because it supervenes upon the upshot of their deliberations" (Richardson, 2018, p. 49)

reasoning, critics argue that it risks oversimplifying moral deliberation, reducing the richness of diversity and failing to adequately account for the complexity of individual autonomy and human experiences.

Pluralists, on the other hand, argue that there is no single fundamental value and that genuine moral conflicts may be irresolvable because different values are incommensurable. They claim that multiple, sometimes competing, values exist, and there may be no objective way to rank them. Pluralists particularly grounded in empirical research acknowledge the possibility of evolutionary core norms that promote coordination for mutual benefit (Gibbard, 1990; Tomasello, 2009), but argues that even if such fundamental norms exist, different societies may develop distinct equilibria in interpreting and prioritizing them, leading to persistent moral incommensurability of values. Moral reasoning, therefore, involves balancing conflicting values rather than seeking a single, definitive answer. Communitarianism for example suggests that different cultures, traditions and shared narratives hold fundamentally distinct and equally valid moral values that are incommensurable (MacIntyre, 1981). While pluralism preserves moral diversity and respects the lived realities of different communities and traditions, critics argue that it can lead to relativism, making it difficult to justify moral criticism across cultures, and moral paralysis over moral dilemmas where values come into direct conflict.

Even though some scholars argue that constructivists endorse a form of value monism, constructivists explicitly align themselves with value pluralism. Rawls famously started his political analysis from the "fact of reasonable pluralism" (Rawls, 1993, pp. 10–14) the idea that even under fair conditions, rational individuals will inevitably hold diverse and often conflicting moral, philosophical, and religious doctrines. While constructivist scholars have expanded, refined, and diverged from Rawls' original position in various ways, they have largely maintained this fundamental commitment to pluralism (Bagnoli, 2018b, p. 154). The misconception that constructivists are monists arises from misinterpreting two intertwined questions: the *source* of value and the *structure* of valuing. Constructivists maintain that Reason is the source of value, but this does not imply that all values are reduced to this singular (source of) value. While Kantian constructivists do posit a structural principle governing all rational activity, this principle is not itself a value but rather the structure through which multiple, incommensurable values are

generated. Kantian constructivism only imposes formal constraints on the way we reason and constitute ourselves but allows for a plurality of values.

Yet, this still leaves open the question of *what kind* of pluralism constructivists endorse. Scholars commonly differentiate between weak incommensurability and strong incommensurability (Bagnoli, 2018b, p. 155). Weak incommensurability suggests that there is no universal scale on which all values can be compared, meaning that while some values may be aligned or weighed against one another, not all values can be subsumed under a single evaluative framework (Wiggins, 1987a). Strong incommensurability, however, posits that any two specific values lack a common metric entirely, making direct comparison impossible. This stronger form implies a deeper pluralism, where values resist ranking or reconciliation by any single measure (Anderson, 1995).

Here it is useful to distinguish between different constructivist approaches and analyze them closer. In fact, certain constructivist scholars are often associated—albeit mistakenly, in my analysis—with a form of commensurable monism. For instance, Forst has been linked to value monism, not only because he places the principle of justification at the foundation of his philosophical system,[345] but also because he seems to explicitly state that justice is not just one value among others but the overriding one. In his words: "It is important to stress here, first of all, that justice is not one "value" among others-like freedom, equality, and so on-but is the principle used to determine which freedoms and forms of equality are legitimate. Justice, which is grounded on the principle of justification, is the first and overriding virtue in political contexts. […] it is important to note that this "monistic" nature of justice allows it to open up to the pluralism of specific aspects of justice (e.g., need and desert) and the uniqueness of different spheres of distribution, in which particular goods (e.g., health care, education, and public offices and positions) are distributed according to particular criteria that also have a socially relative character" (Forst, 2012, pp. 7–8).

And yet, in more recent works, Forst also specifies that "the version of constructivism I advocate does not assume that there are no values other than those constructed in this way. There

[345] As I have previously argued, positing a structuring principle of rationality does not necessarily imply monism, especially in debates on commensurability. A structuring principle does not function as a preeminent value among others; rather, it provides the framework within which deliberation about values takes place.

are various sources of normativity, some of which are specific (particular communities of shared values), some general (a universalistic religion, for example). And in the sphere of morality as well, it is not necessary to dispute the reality of certain values or reasons in principle. My approach remains agnostic regarding these questions. The crucial point is that, in the context of morality (not in that of the good in general), the principles and criteria of practical reason are understood as transcendentally and pragmatically" (Forst, 2024, p. 8). Forst, therefore, endorses a form of weak incommensurability. His position is not monistic, as conflicts of values are not resolved by appealing to a supreme, overriding value like justice— although it holds a primary role among values—but rather through the principles of reciprocity and generality, which are constitutive of the principle of justification. At the same time, his view does not amount to strong incommensurability, since he acknowledges that some value conflicts can be resolved by critically examining and exposing the justifications behind them, distinguishing between legitimate reasoning and ideological distortions (Forst, 2024, pp. 51–52).[346]

Without providing a detailed analysis of other constructivists —a task that would require extensive independent work— I will simply note that there is significant evidence suggesting most constructivists subscribe to a form of weak incommensurability (Bagnoli, 2018b, p. 156). This means that while there is no overarching fundamental value that can always resolve conflicts, certain non-reducible values can still play a primary role, provided their justification is properly articulated. Rawls, for example, argued that the basic liberties of individuals (e.g., freedom of speech) have "lexical priority" over considerations of economic efficiency or general welfare (Rawls, 1971, p. 42).[347]

[346] "There are a number of concepts in our moral vocabulary that I call "normatively dependent concepts," as they only gain normative substance by being connected to other normative principles or values. Toleration, solidarity, legitimacy, and trust are examples of such concepts. They are often mistakenly seen as values in themselves, but actually they are not; for example, solidarity can be a good or a bad thing depending on the justifications for it." (Forst, 2024, pp. 51–52). For discussions on the specific concepts, see Forst on toleration (2013, p. 32), solidarity (2024, Chapter 4), legitimacy (2017a, Chapter 8), and trust (2022, p. 2).

[347] Importantly, Rawls here is dealing within the specific domain of distributive justice, a distinct area of political philosophy concerned with the fair distribution of rights, duties, and social goods. So what he is implying is that in this practical context, the basic liberties cannot be overridden by arguments appealing to greater welfare or efficiency. "Lexical priority" means that between two kinds of values—say, a and b—if a has lexical priority, a must always prevail over b, regardless of the consequences. However, the fact some values in some specific context can have lexical priority over other does not entail that the constructivist framework is committed to any form of hierarchical strong incommensurability.

The idea that some values are higher and beyond trade-offs allows us to reconnect with our broader discussion on rights. A pluralist approach that endorses weak incommensurability explains the language of rights by conceptualizing them as protections for distinct, non-reducible values, serving as a mediating framework in conflicts between competing values (e.g., liberty vs. security, equality vs. efficiency, etc.); since no single value overrides all others, rights are necessary precisely because they establish moral boundaries that remain respected even in cases of deep disagreement. This is precisely what Dworkin meant by saying that rights are trumps: they not only shield individuals from government actions that might infringe upon their moral and legal entitlements in pursuit of broader social or economic goals, but they also provide individuals with a protective standing in interactions with others—especially in contexts where different agents endorse competing values or fail to recognize the priority of certain rights in specific situations. Rights in this why function as a legitimate demand of protection for certain sort of considerations that "overrides any other sort of consideration" (Dworkin, 1977, p. xi).

2.4 BACK TO RIGHTS, AND BEYOND

The main aim of this chapter was to understand how constructivists develop a normative conception of rights, by analyzing the conceptions of normativity they endorse and how this shapes other normative concepts like reasons and values. By conceiving normativity as a form of self-legislation, constructivists directly link it to agency, specifically autonomous agency. In the Kantian sense, autonomy means that individuals are not merely passive recipients of moral norms but actively participate in shaping them through rational reflection. Autonomous moral agents endorse norms not because they are externally imposed but because these norms can be justified through their own reasoning. By explaining the source of normativity, constructivists clarify how moral principles become normatively binding for individuals (Bagnoli, 2013b, p. 120).

Reasoning, therefore, is the process through which norms are constructed—not merely through deduction or by deriving them from fixed first principles, but through a practical and dynamic process. While there is a constitutive element that grounds the legitimacy of this constructive process, reasoning also incorporates an intersubjective (Forst, 2002, pp. 242–245;

269–273) and transformative (Bagnoli, 2016, pp. 1229–1230, 1244–1245) component of deliberating with other agents. Reasons for actions are the end of this practical deliberation process, not the starting point as some moral realists argue. The strength of constructivism lies precisely in its ability to account for the objectivity of reasons without relying on a metaphysically burdensome ontology, but only on the formal structure of how we reason, namely universalizability. To say that I have a reason to keep a promise means that, by reflecting on the conditions of the act of promising itself, I recognize that I cannot arbitrarily justify freeing myself from the commitment I made. The very practice of promising presupposes the expectation of reliability: if I were to break my promise without justification, I would undermine the very meaning and function of promising as a practice whit is correlated norm. This is not because an external, independent moral fact compels me to keep my promise, but because, through rational deliberation, I recognize that my commitment entails obligations that others, under the same conditions, could also recognize and hold me accountable for.

Similarly, the concept of values is not the beginning of the deliberation process, but the end of it, although it is used more broadly by constructivists, encompassing concepts that vary from actions and institutions, to entity, state of affairs and traits (Bagnoli, 2018b, p. 148; Korsgaard, 1983, p. 169). To say that friendship is valuable is not primarily to claim that it should be promoted or that a world with friendship is necessarily better. Rather, it is to recognize that we have good reasons to value friendship itself. Certain intrinsic characteristics of friendship lead us to this conclusion; these qualities provide the empirical foundation upon which we construct reasons for valuing and acting toward this specific relationship. However, these qualities do not have an independent ontological status—they are simply the raw empirical material we use to orient ourselves meaningfully in the world. Thus, to value something is to have reasons toward it or, to use Scanlon's formulation, having reasons for holding a positive attitude toward it (Scanlon, 1998, p. 95). Friendship is not a value, which would seems to endorse a foundationalist stance (Sensen, 2011, p. 30), but has value because we have genuine reasons to seek and nurture it. Constructivists, in this regard, are not far from Scanlon's buck-passing account of value; however, they diverge

from him since the reason through which we specify what is choiceworthy are not primitive concepts; reasons are also constructed through reasoning, because Reason precedes reasons.[348]

Furthermore, constructivists are not committed to the claim that reasons precede values (Scanlon, 1998), or vice-versa (McDowell, 2002),[349] only that Reason precedes reasons and values. In this view, both reasons and values are constructed through the activity of Reason, rather than existing as independent, pre-given entities.[350] This means that when we specify a final end (value) like friendship we are reasoning on why it is choiceworthy; in doing so it is possible to deliberate directly on this value, without the mediation of specific reasons for action (or at least without the need for a strict distinction between these concepts in the process of deliberation).

This helps clarifies why rights, and specifically human rights, contrary to what is stated and even ratified in documents, should not be conceptualized as self-evident values, mind-independent entities that give us reason to pursue them. Instead, similarly to how reasons and values "are the results rather than the starting points of practical reasoning" (Bagnoli, 2013a, p. 21), for constructivists rights are the outcome of practical reasoning and political deliberation, rather than the starting points they are often assumed to be. Rights are necessary tools to balance the fact of value pluralism, because they function as a demand of protection for certain considerations and concerns, but they are truly normative, for all rational beings, when this demand is legitimate.[351]

---

[348] Here of course we are referring only to the reasons that are not constitutive of reasoning but constructed from it. This is one of the reasons to prefer using 'constitutive norms' over 'constitutive reasons', to leave the concept of reasons only to the construction part and avoid ambiguity.

[349] McDowell explicitly argues that practical rationality entails the ability to be responsive to reasons for action that stem from value. Drawing parallels with aesthetics (McDowell, 2002, pp. 112–120) he contends that when we recognize an artwork's aesthetic value—a (secondary) property of the world, albeit perceived through our sensibility—this value itself provides a reason to contemplate, preserve, or share it with others. Similarly, in the moral domain, recognizing the value of justice or compassion can provide us with reasons to act justly or compassionately. Moral value is not merely a subjective reaction but a feature of the world (or of human interactions) that can both motivate and justify our actions, thereby providing reasons for them.

[350] To use a more appropriate vocabulary, we could say that both of them are acts of Reason, but while a value is a concept, a reason is a proposition or judgment. In this sense, most constructivists would argue that specifying some concepts of values requires making judgments about why they are valuable, thereby establishing a certain priority of reasons over values. Still, this priority differs significantly from Scanlon's view, in which reasons are taken as primitive concepts (Scanlon, 1998, p. 1).

[351] While this might seem to shift the problem from pluralism about competing values to pluralism about competing rights—cases where two or more rights-based claims conflict—conceiving both as outcomes of the

What makes the demand legitimate and the right normative stems from autonomous reasoning about why it should be so. If we cannot reasonably justify it to the agents responsible for fulfilling the corresponding duties (the supposed duty-holders), or if the burden placed on them is too high, then the claim is illegitimate. Similarly, we must be able to justify to the subject demanding the right (the supposed right-holder) why their request is not valid. This justification occurs through deliberation, both internally—via the constitutive norms of reasoning—and externally—through the moral norms revised and accepted by the moral community. Just as the specification and ratification of a moral norm is an open-ended process, so too is the specification of a moral right.

In this sense, we can reformulate the provisional definition of rights we previously offered (legitimate specificatory demands) by characterizing them as *reasonable specificatory demands*; this allows us to clarify what makes such demands legitimate and normative from a constructivist perspective, namely our Reason. The revised definition applies to both moral and legal rights: the difference lies in the kinds of reasons that justify the legitimacy of the demand. For moral rights, the reasons are typically general in nature, while for legal rights, they may be more contingent, depending on institutional or contextual factors. The gulf between moral rights and legal rights is, therefore, not a matter of nature but of application, particularly in how the "social division of moral labor" (Richardson, 2018, pp. 127–128, 158) allows for the assignment of specific duties to individuals or groups to ensure the institutional fulfillment of those rights. This process not only

deliberative process helps avoid this dilemma. To illustrate, consider how constructivists resolve the frequent conflict between religious rights and gender equality rights, especially in contexts such as dress codes, employment, and education. For example, the controversy over Muslim women's and girls' wearing of the hijab, has sparked intense debate in many liberal-democratic societies, particularly regarding teachers' headscarves. The two opposing views are the following: one sees the hijab as a symbol of intolerance and female oppression incompatible with liberal values, while the other advocates for tolerance and equal rights unless there is clear harm. This tension reflects deeper disagreements about toleration itself, revealing how some societies still treat minority rights as conditional privileges rather than fundamental protections (Forst, 2012, p. 150). After careful consideration, thus, it becomes clear that in this specific case the apparent conflict between the two rights is only superficial, provided that wearing the hijab does not cause direct harm to others but merely constitute an expression of religious freedom. Whether the ideological element of religion influences the freedom of this decision remains an important question. While it is important to consider whether the ideological influence of religion might affect the authenticity of the choice to wear the hijab, such influence cannot be presumed by default, but must be verified case by case and agent by agent. From a constructivist perspective, the autonomy of the individual agent—and their capacity to engage in practical reasoning and self-legislation—must be taken as a starting point. Therefore, rights may be understood as *trumps*, to borrow Dworkin's terminology, in the sense that they offer protections that cannot be easily overridden for the sake of general welfare or utility (Dworkin, 1977, p. xi). However, unlike in Dworkin's account, their normative force does not stem from any intrinsic moral authority (Sensen, 2011, p. 51), but from the fact that this authority can be justified through principles of practical reasoning that all individuals can reasonably accept (Rawls, 1997, p. 801) and the role they play within a constructed and mutually justifiable normative order (Forst, 2024, pp. 3–4).

designates the institutional duty-holder but also defines the categories of beneficiaries who enjoy the protections or advantages of that right. However, the recognition of rights still presupposes active reasoning, as only by assessing the specific material conditions at hand can we determine which individuals have legitimate claims to certain protections. For example, workers have a legitimate claim to safe working conditions and fair wages, which justifies the establishment of workers' rights as a recognized category.[352]

But what about the category of human rights? Does the beneficiary of these rights encompass all of humanity as a single group? If so, where do human rights derive their legitimacy, namely, how can we justify that all human beings, by virtue of their humanity alone, hold specific reasonable claims that must be respected by other agents and institutions? These delicate questions will be the focus of the next chapter.

[352] "Finally, simplifying and idealizing, there is the possibility that society contains institutions that realize, as far as is reasonably possible, the human right to employment security. So long as that condition lasts, the unspecific address of this right will recede in importance, being for practical purposes largely superseded by its being specifically addressed to the relevant institutions (the employment security agency, for instance) and their officers. As the example of employment security illustrates, the contingent establishment or evolution of social institutions can give rise to the existence of people who, by virtue of their role or office, have special duties to see to it that a given right is fulfilled or respected— duties that are ultimately grounded by reference to that right." (Richardson, 2018, p. 98).

CHAPTER 3

THE IDEA OF HUMAN BEING

The insight that all human beings have acquired the "standing" as beings to whom justification is owed insofar as they are human beings alone, is not the result of philosophical deductions and stipulations alone but the achievement of hundreds of years of moral and political struggle. Philosophy has the task of clarifying the lessons of these struggles as well as giving the most rationally justifiable defense of their ideals.

(Benhabib, 2015, pp. 784–785).

We have examined how rights in general are conceptualized within a constructivist framework (Chapter 1) and analyzed the specific type of normativity they carry and their relations to other normative concepts (Chapter 2). We concluded that, within this framework, a right qualifies as a legitimate demand only if it can be reasonably justified to those responsible for fulfilling the corresponding duties. This justification is both internal—via the constitutive norms of Reason—and external—through the public deliberation of moral norms by the moral community. The process of specification can be carried out reciprocally, by directly addressing either the duty-holder or the right-holder, or generally, by reasoning from a shared moral standpoint and imagining how any agent would be treated under similar conditions; when circumstances allow, these two modes of specification can (and should) complement one another.

In this chapter, we will finally introduce the beneficiary aspect of what makes some rights human rights, as well as the supposed foundation that grants the main justificatory force to their

enactment, seeking to explain how a constructivist theory should address this dimension.[353] If we provided a constructivist definition of right as a reasonable specificatory demands, can we simply incorporate human beings in this definition to achieve a satisfying constructivist definition of human rights? Forst for example define human rights as "fundamental rights every human beings can claim as human beings" (Forst, 1999, p. 43). Do we really need an extensive clarification of the concept of human beings to gain a clearer understanding than this? As we discussed at the beginning of Part I, the concept of a human being appears to be already clear and self-evident, suggesting that further definition may be unnecessary, or we might simply refer to the established scientific definition.[354]

This sidesteps the central question: what specific characteristic of being human confers an entitlement to fundamental rights?[355] Does the concept of dignity sufficiently address this question, or does it give rise to problems of regression or circularity? In what way does this characteristic have justificatory power, and why does the same not apply to other non-human animals? [356] To

[353] This is a crucial task, especially considering that constructivism and foundationalism are typically considered as antithetical (Grube, 1998; Bagnoli, 2013a). However, as will become clearer in this and in the next chapters, grounding human rights does not necessarily lead to foundationalism.

[354] Criticisms can be raised against this proposed simplicity, for example, emphasizing how the mere biological definition fails to truly capture what makes the human species unique, distinguishing it radically from other types of living beings. Therefore, not only because they are holders of rights, but also because they possess value (an objection we will return on section 3.3) Another more substantial criticism refers instead to the debate on the concept of species itself, which is less clear than is often thought. For an examination of the ontological and epistemological problems related to species taxonomy, see, for example, Ereshefsky's treatment (2022).

[355] Forst as well further specifies his definition by describing human beings as justificatory beings. He writes: "Philosophy has defined human beings in numerous ways: as beings that are endowed with reason (animal rationale) and equipped with the unique capacity for language (zoon logon echon), that are also finite and limited, "flawed beings", and last but not least as social (animal sociale) and political beings (zoon politikon). In my view, what emerges from the combination of these definitions is the image of human beings as justificatory beings. They not only have the ability to justify or take responsibility for their beliefs and actions by giving reasons to others, but in certain contexts they see this as a duty and expect that others will do the same." (Forst, 2012, p. 1)

[356] There is a further question of whether the individual is the sole bearer of rights, or whether rights also extend to other entities, such as families, nations, and ethnic, religious, and linguistic groups. In fact, this is already the case for corporations and companies, which have specific rights despite being entities comprised of private citizens. In this chapter, I will not delve into the details of this issue, assuming that such rights should be considered at a later stage. Defining the rights of an individual inherently establishes normative relations with other bearers of the same status, thereby addressing the social and collective aspects (Richardson 2018). See also footnote n. 483. For an alternative approach that does not initiate with the rights of subjects, one can refer to Miguel Vatter, who expands the question to encompass the rights held by all natural entities as integral components of the Earth's bio-organism (Vatter, 2020).

explore these queries, we will start by scrutinizing how the concept of dignity, a concept often left undefined and simply assumed to be self-evident,[357] is employed in the human rights discourse.

3.1 THE CONCEPT OF DIGNITY IN THE HUMAN RIGHT DEBATE

In the Preamble to the Universal Declaration of Human Rights, dignity holds a central place: "Whereas recognition of the inherent dignity of all members of the human family and of their rights, equal and inalienable, constitutes the foundation of freedom, justice and peace in the world [...] the General Assembly proclaims this Universal Declaration of Human Rights as a common ideal to be attained by all peoples and all nations, to the end that every individual and every organ of society, keeping this Declaration constantly in mind, shall strive by teaching and education to promote respect for these rights and freedoms and by progressive measures, national and international, to secure their universal and effective recognition and respect, both among the peoples of Member States themselves and among the peoples of territories under their jurisdiction".

This emphasis on the foundational manifest role of dignity is reiterated in the simplicity and clarity of Article 1: "All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and must act towards one another in a spirit of brotherhood". This preamble fails to offer a clear understanding of what dignity is or the role it plays, as it lacks a specific definition. Instead, dignity is merely "recognized" as "inherent" and positioned as a foundation left to intuitive understanding[358]. When it has been invoked in concrete situations, it

[357] The literary critics Mark Greif has argued that discussions around dignity are always "cant discourse", in the sense that they are not meant to be questioned or refined through disagreement. This is due to the fact that this "cant terms, like dignity, originated in a theoretical framework that is no longer alive to its users, so they create an atmosphere of profundity and mystery in a purely symbolic way" (Greif, 2015, 11). Theunissen agrees and expand on this, arguing that discussion of dignity "tend to involve a too-easy appropriation of Kantian terminology with inadequate attention to foundational questions, including the question of what value is. The term becomes harmful because it becomes empty, and in becoming empty it drains the discussion." (Theunissen, 2020b, p. 5)

[358] It has been argued that avoiding a rich philosophical account of dignity is intentional, as there's a fear that such an account might appeal to 'comprehensive doctrines' (to use Rawls's term) and potentially cause division. Andrea Sangiovanni is convinced that there is a common agreement that providing a deeper grounds for human rights will limit their appeal, and this is way "philosophers of human rights have wanted to remain ecumenical, following in the spirit of the Universal Declaration and the various human rights covenants, practices, movements, organizations that have followed in its wake" (Sangiovanni, 2019, p. 250). There are two primary reasons why we should invert this trend. First, not discussing the reasons behind our belief in human dignity won't convince those who think it is

has been generally assumed that a violation of human dignity can be recognized even if the abstract term cannot be defined[359]. The concept of dignity, therefore, seems to manifest all the problematic features we have identified in a jusnaturalist framework with foundationalist tendencies. As Simmons noted, the idea underlying these types of approach can produce unfruitful results in both cases: "The idea of human dignity is not best understood as the proper foundation of a theory of human rights at all, or even as something (more basic but not foundational) from which human rights can be derived. […] It might be a mistake to suppose that a theory of human rights is a better or more fully justified theory when it is made to rest on such deep and naturalistic foundations as an account of human dignity or a theory of natural rights" (Simmons, 2015, p. 139)[360].

Many philosophers are rather skeptical that the concept of dignity serves any specific function not already fulfilled by other concepts (such as value, autonomy, or respect). As Adam Etison noted, a good theory of dignity need to pick out a "meaningful distinct set of concerns" if it is to justify using the term at all (Etinson, 2020, p. 354). We can generally identify three distinct trends in the approach philosophers have toward the concept of dignity:

---

acceptable to treat some individuals with less respect. Without understanding why we possess dignity, we're likely to struggle with determining its implications, especially in clarifying why and when human rights override strong normative considerations like maximizing general utility, advancing certain ideologies, or protecting traditional practices. Secondly, neglecting to explore the concept of human dignity leaves us directionless on crucial questions related to human rights theory. As we have mentioned, for instance, it is improbable that we will comprehend which rights should be considered human rights without first understanding the rationale behind their existence. Answering this and similar questions becomes challenging without a clear grasp of what dignity entails (Caranti, 2019a, p. 62).

[359] Numerous judicial decisions have invoked the concept of dignity without making an effort to provide a general definition. For instance, reference can be made to the judgment of the European Court of Human Rights in Lawless v. Ireland (Merits), 1 EHRR 15, 39 (judgment of July 1, 1961). Similarly, the courts of West Germany have issued several noteworthy decisions applying the constitutional provision on human dignity. But although in some situations an abstract definition is not needed "it is not entirely satisfying to accept the idea that human dignity cannot be defined or analyzed in general terms. Without a reasonably clear general idea of its meaning, we cannot easily reject a specious use of the concept, nor can we without understanding its meaning draw specific implications for relevant conduct" (Schachter, 1983, p. 849).

[360] Simmons extends this critique to the entire foundationalist problem, not only to the concept of dignity (which is, in fact, partially reintegrated by the author). In the conclusion of his paper, he writes that "we do far better in understanding human rights, I think, to continue to pursue and criticize the arguments of natural rights theories than we do in trying to justify remaining neutral in such debates or in constructing blanket (and necessarily distorting) rejections of all "naturalistic" theories of human rights. In the end, I suppose, we might all come to agree that the foundational arguments of natural rights theory are in some way fatally flawed, or that they cannot constitute the basis of a satisfactory theory of human rights. But this has not yet been convincingly shown. The game, I think, is still afoot" (Simmons, 2015, p. 152).

1) Remove it in favor of other terms

2) Maintain it with all its theoretical roots.

3) Reconceptualize it by providing new definitions that specify a new meaning.

Clearly, the boundaries between these approaches are blurred, and even those who seek to reconceptualize the term often draws, more or less explicitly, to multiple historical currents at the expense of others.[361] I will begin by introducing scholars who explicitly advocate for the removal of the term 'dignity' from human rights discourse. Following this, I will explore the more nuanced distinction between scholars who seek to preserve one of the original meanings of dignity and those who aim to reframe or reconceptualize it. As we will see, philosophers often associated with constructivism are found in both the second and third groups. However, I will argue that the most original contributions come from the latter group, where the concept undergoes significant reinterpretation.

3.1.1 Eliminating Dignity: Macklin, Pinker and Singer

Critics of the concept of dignity has largely focused on bioethical concerns, stemming from the fact that references to dignity in major international human rights instruments, such as the United Nations' Universal Declaration of Human Rights, do not specifically address medical

[361] This issue arises with any taxonomy that attempts to encompass and synthesize the diverse approaches to a broad and complex subject such as dignity. Oliver Sensen, for instance, highlights a widely recognized distinction between three paradigms of dignity (Sensen, 2011, pp. 147–173). First, the 1) *Traditional (or Aristocratic) Status-Based Paradigm,* where dignity was associated with rank, honor, and social status—a privilege reserved for those of higher standing, such as nobility or specific professions. Second, the 2) *Kantian Inherent Worth Paradigm*, which holds that all human beings possess intrinsic worth simply by virtue of being rational agents capable of moral reasoning. Finally, the 3) *Contemporary Human Rights Foundational Paradigm*, in which dignity is interpreted as a foundational principle that underpins universal human rights, serving as the basis for fundamental entitlements such as autonomy, equality, and freedom. argue that, in fact, Kant's use of dignity is more closely aligned with the Aristocratic tradition and that Kant does not, as often assumed, endorse an inherent worth value-concept of dignity. Thus, Sensen reduces the paradigms to two: the traditional status-based notion of dignity and the contemporary foundationalist notion. While this historical bipartition is useful, it risks excluding certain contemporary perspectives that cannot easily be subsumed under a form of value foundationalism. Sensen acknowledges this potential oversimplification since he also further elaborates on the traditional approach, dividing it into four distinct theses: 1) Dignity is not a value; 2) It is a dual-faceted concept, capable of both realization and degradation; 3) It does not inherently confer rights; and 4) It emphasizes duties to oneself, particularly in the context of self-perfection (Sensen, 2011, pp. 161–164). We will return to Sensen's interpretation of Kant in paragraph 3.2.1

treatment or research.[362] In the U.S. for example, much of this critique has been directed at the use of 'dignity' by the President's Council on Bioethics, an advisory body established by President George W. Bush on November 28, 2001.[363] Ruth Macklin in particular argued in her essay "Dignity Is a Useless Concept" that appeals to dignity are often redundant, merely echoing the established principles of medical ethics, like respect for autonomy, the necessity of obtaining informed consent, the protection of confidentiality, and the imperative to prevent discrimination and abusive practices (Macklin, 2003). Since in most treatises the term 'dignity' is essentially used as a synonym for capacities such as rational thought and autonomous action, Macklin argues that it contributes little to the discussion and that, consequently, the term can be regarded as superfluous and potentially omitted (Macklin, 2003, p. 1419).

Steven Pinker and Peter Singer take the critique of dignity even further, not only arguing that the term is redundant or lacks conceptual clarity, but also claiming it serves an illogical rhetoric, often smuggled into public discourse by religious individuals or institutions who oppose moral autonomy and seek to obstruct progress in cutting-edge medical research. According to Pinker, religious leaders have consistently attempted to stifle the spread of empathy-promoting literature and have historically resisted reforms such as the abolition of cruel punishments and slavery. In Pinker's view, religion is deeply intertwined with humanity's "inner demons", which manifest in the desire to dominate others, seek revenge, and pursue utopian ideals that justify violence in the name of some imagined greater good (Pinker, 2012).[364] Because religion is inherently oppressive, Pinker warns that appeals to dignity by religious figures are often a tactic to undermine scientific inquiry and maintain control over others. Therefore, he suggests we should be highly skeptical of the use of the concept of dignity and focusing instead on cultivating secular virtues such as empathy, self-control, and respect for reason.

Singer also takes a skeptical stance toward traditional views on human dignity, particularly the idea that humans possess inherent and unique moral worth simply by virtue of being human,

[362] A leading exception is the Council of Europe's convention for the protection of human rights and dignity of the human being with regard to the application of biology and medicine

[363] Executive Order 13237 tasked the Council with several key objectives, to address ethical questions related to areas such as embryonic and stem cell research, assisted reproduction, cloning, applications of knowledge from human genetics or neuroscience, and end-of-life issues.

[364] See especially chapter 8, called precisely "Inner Demons".

and he highlights the religious influences that have shaped this conceptualization (Singer, 2011, p. 239).[365] Since values are not metaphysical entities "out there" in the universe waiting to be discovered, reason's role in ethics is precisely to recognize universal principles that transcend our narrow self-interest: equal consideration should be extended to all sentient beings, not exclusively to our own species. Singer, however, formulates the challenges to human exceptionalism in strictly utilitarian terms: the capacity for suffering and enjoyment—not mere membership in the human species—should determine moral consideration. Consequently, many animals, as sentient beings capable of suffering and experiencing well-being, deserve moral consideration equal to that of humans (Singer, 1976).

Beyond the classic challenges associated with utilitarianism—such as the difficulty of accurately calculating actions to maximize overall good and the risk of instrumentalizing individuals—Singer's and other eliminativist approaches to dignity may risk undermining the foundation of human rights, which seems to rely on the premise that all humans deserve equal respect simply by virtue of being human. By suggesting that moral worth depends on cognitive and experiential capacities rather than species membership, Singer's approach could weaken arguments for universal human rights protections, making them subject to varying interpretations.[366] This shift could reduce protections for vulnerable groups, such as elderly individuals with dementia or those with severe cognitive disabilities, as it may imply that their lives hold less moral worth than those of fully functioning adults.

3.1.2 Preserving Dignity: Waldron's Reaffirmation of a Status-Concept

Those who seek to preserve the term "dignity" often appeal to at least one of its established usages in everyday language, which have been transmitted through historical and cultural continuity. For example, Teresa Iglesias, starting precisely from the definitions found in dictionaries, distinguishes between a universal and a restrictive meaning of the term dignity,

[365] "Western attitudes to nature grow out of a blend of those of the Hebrew people, as represented in the early books of the Bible, and the philosophy of ancient Greece, particularly that of Aristotle. In contrast to some other ancient traditions, for example those of India, both the Hebrew and the Greek traditions put humans at the centre of the moral universe. […] [H]umans are not merely of central moral significance, they constitute the entirety of the morally significance features of this world." (Singer, 2011, p. 239).

[366] This is in accusation that can sensibly be asked also for some "rationalist elitists" which sometimes also looks to Kant as a point of reference. We will return to this point in 3.4.

associating the restrictive use with classical Roman culture and the universal use with the notions of intrinsic human value that emerged in Jewish ethics and theology (Iglesias, 2001, p. 120).[367] However, although "the meaning of dignity has historically been marked, until now, by a tension between its universal and restrictive meanings", what has happened is that "historically, the restrictive Roman meaning of *dignitas* assigned to office and rank, and used as a discriminatory legal measure, has begun to be used with a new meaning of universal scope that captures the equal value of all." (Iglesias, 2001, p. 122). In this sense, then, Iglesias maintains the significance believed to be original to the term while explaining its evolution and intertwining meanings.

Jeremy Waldron, one of the most influential thinkers on this topic, has also sought to reclaim the term, drawing from its theoretical Roman roots while simultaneously implementing a form of re-semantization compared to its current usage. In opposition to the prevalent contemporary interpretation of dignity as a value, Waldron contends that it should be understood as a status. This re-semantization seeks to re-establish a connection with the Roman use of the term *dignitas*, which "embodied the idea of the honor, the privileges, and the deference due to rank or office" (Waldron, 2012, p. 30).[368] The modern notion of human dignity seems at odds with this ranks connection, as highlighted by Iglesias, but according to Waldron, this change happened harmonically, with its acmè on the late-eighteenth-century romantic poetry.[369] This entails an "upwards equalization of rank, so that we now try to accord to every human being something of

[367] "Consulting the dictionary, we can find that the term 'dignity' connotes 'superiority' and 'decorum' related to it, in two basic senses. One refers to superiority of role or rank, of office, of excellence, of power, etc., which can pertain only to some human beings... The other refers to the superiority of the inherent worth of every human being which is independent of the external conditions of office, rank, etc., and which affects everyone. In this universal sense, the word "dignity" captures the specific way of being of the human being as a human being. The latter meaning, therefore, has a universal and unconditional meaning, in contrast to the former which is restrictive and role-determined." (Iglesias, 2001, p. 120).

[368] Even for Waldron, the use of the word dignity (and especially the Catholic use of it), could be replaced by other terms such as value or "sacred worth" (Waldron, 2012, p. 28). However, ultimately, although he does not want to diminish the rift between Roman-Greek and Judeo-Christian ideas, he still believes that the connotation of rank status remained and that "what happened was that it was transvalued rather than suppressed" (Waldron, 2012, p. 31).

[369] The impulse to universalize the concept of dignity to encompass all human beings was already evident in the Roman period, particularly in Chapter XXX, Book I of 'De Officiis' by Cicero, which according to Viktor Pöschl, is precisely the earliest known instance where the Latin term 'dignitas' is applied to all human beings (Pöschl, 1969, pp. 38–41). Although Greek philosophy often echoed a similar pattern of thought regarding the value and status of individuals, there was no direct equivalent to the Latin term 'dignitas' in the Greek lexicon (Pöschl, 1969, p. 9). This historical reconstruction of the term is shared by many academics, from Baker (1947) to Forst (2011).

the dignity, rank, and expectation of respect that was formerly accorded to nobility." (Waldron, 2012, p. 33).

If Waldron is right that dignity is a status concept and not a value concept, then we should conceive of dignity, akin to other statuses, as "a particular 'package' of rights and responsibilities" that an individual possesses by virtue of their position or situation (in this case, by virtue of being human); in this sense, human dignity becomes both a shorthand term for the list of human rights we find in international practice, but also includes "the underlying idea that informs the status in question" (Simmons, 2015, p. 141). This idea, as mentioned, entails the assignment of a high status in virtue of merely being human, enshrining a profound link between dignity and egalitarianism.[370]

Most constructivists, as we will explore, generally agree that if we are to retain the term "dignity" it should be understood as a status-based concept. However, they differ in their emphasis on the historical origins that underpin this conceptualization. For instance, Ronald Dworkin (2011), which also acknowledges the inextricable interpretative process involved in adhering to a specific tradition,[371] argues that dignity encompasses two core values deeply rooted in Western political theory: the notion that every human life possesses intrinsic value and the idea that individuals bear a special responsibility for the course of their own lives.[372] It is precisely with regard to the first core value—that human life possesses intrinsic value—that some constructivists distance themselves. In the following pages and throughout the rest of this chapter, I will argue that the most compelling constructivist approaches are those that prioritize the second core value identified by Dworkin. These approaches emphasize individual responsibility for one's life rather than relying on the notion of intrinsic value; by doing so, they offer a conception of dignity as a status that is not necessarily grounded in intrinsic worth but rather in the dynamic and relational aspects of personal and social responsibility

[370] James Whitman defined it as "an extension of formerly high-status treatment to all sectors of the population (Whitman, 2005, p. 108).

[371] "[D]ignity, like so many concepts that figure in my long argument, is an interpretative concept" (Dworkin 2011, p.200).

[372] These moral principles of dignity further serve three distinct functions. First, they offer guidance on personal ethics, outlining what is required for living a well-rounded life. Second, they clarify the rights individuals possess vis-à-vis their political communities. Third, they articulate the moral obligations we have toward others (Dworkin, 2011, pp. 193–217).

3.1.3 Reconceptualizing Dignity: Constructivist Redefinitions

The literature on the concept of dignity is extensive and frequently intersects with various other topics, making it difficult to detect when a philosopher introduces a new definition of the term. This complexity is also due to the tendency of some authors to hide their theoretical innovations with established predecessors to lend legitimacy to their ideas. James Griffin, a prominent scholar in the human right debate, can be considered one of these cases. Although he explicitly claims to recover the concept of dignity as understood by Pico della Mirandola, Griffin's work more accurately represents a process of re-semantization, blending the term with the concepts of normative agency and personhood: dignity, in Griffin, is the human capacity to "be that which he wills" (Griffin, 2008, p. 31). Unlike other creatures, human beings do not have a nature that is determined in advance: they are self-creators. The concept of self-creation here is intricately linked to autonomy because it unfolds through choices guided by reason. The sort of dignity relevant to human rights is thus that of a highly prized status: that we are normative agent (Griffin, 2008, p. 152). In this sense, human rights are both derived and a protection of dignity. Affirming that human rights are derived from our dignity, and defining dignity as normative agency, an underlying status that any agent has, implies also that human rights have to be considered as a protection of our normative agency.

Moreover, Griffin affirms that this notion of dignity, which he also interchangeably calls personhood, can be unpacked in other three more substantial concepts: that of autonomy, minimal provision and liberty. "One can break down the notion of personhood into clearer components by breaking down the notion of [normative] agency. To be [a normative] agent in the fullest sense of which we are capable, one must (first) choose one's own path through life—that is, not be dominated or controlled by someone or something else (call it 'autonomy'). And (second) one's choice must be real; one must have at least a certain minimum education and must have at least the minimum provision of resources and capabilities that it takes (call all of this 'minimum provision'). And none of this is any good if someone then blocks one; so (third) others must also not forcibly stop one from pursuing what one sees as a worthwhile life (call this 'liberty')" (Griffin, 2008, p. 133). So, according to Griffin, all human rights are either one or the other of these three

highest-level rights or are derived from them.[373] Griffin explicitly asserts that this re-semantization has the advantage of maintaining a reasonable fit with the idea, prominent in major IHR documents, that human rights are grounded in the dignity of the individual. To adopt the personhood account of human rights is to adopt normative agency as the interpretation of 'the dignity of the human person' when that phrase is used as the ground of human rights (Griffin, 2008, pp. 149–158).

However, from a constructivist perspective, Griffin's approach poses a challenge. Despite centering dignity on the notions of normativity and reasons, Griffin presents a realist theory of human rights: the justification for human rights claims ultimately traces back to the recognition that normative agency is valuable per se. He thinks that, when we properly appreciate the value of normative agency, we understand that it is of great intrinsic value, not just in our own case but wherever it exists. Thus, Buchanan critiques of Griffin affirm that his theory suffers the same indeterminacy that it was supposed to avoid. "The difficulty for Griffin is that […] the idea of being able to exercise reasonably effective or adequate normative agency, where this includes some package of liberties, has just the sort of indeterminacy that Griffin seeks to avoid. How much scope for the exercise of normative agency is enough and how effectively must an individual be able not only to form but also to pursue her conception of a worthwhile life? So far as I can tell, Griffin provides no satisfactory answer to these questions" (Buchanan, 2010, pp. 695–696).

This is a problem that the majority of constructivists will seek to avoid, by focusing not only on the agency of the individual subject, but on the normative intersubjective relationship that agents share. Darwall, for instance, aptly highlights the central role of dignity by distinguishing between two types of respect: appraisal respect and recognition respect. The former is an evaluative attitude, expressing esteem based on specific qualities or actions, while the latter is a second-personal standpoint, implying mutual accountability of persons. According to Darwall, dignity is grounded precisely in this recognition respect and consists in giving due weight to the second-personal authority that others inherently possess as persons.[374] By reconceptualizing

[373] We will return on this subdivision while discussing the list problem in chapter 4.

[374] "When I first distinguished recognition respect from appraisal respect […] I argued that whereas appraisal respect is a form of esteem, recognition respect is something we realize in our treatment of others by regulating our conduct towards them by giving adequate weight to the fact that they are persons. […] We fully respect persons only through relating to them second personally" (Darwall, 2013b, p. 14).

dignity in this way, Darwall also distances himself from Waldron's conception, by affirming that: "dignity is an irreducibly normative rather than a socially constituted status" (Darwall, 2013b, p. 16).[375]

Similarly, one of the most prominent constructivist accounts of dignity in the contemporary debate, that of Rainer Forst, situates dignity at its core in the rational recognition of ourselves and others as beings capable of justification. Forst describes individuals as "justifying beings" (Forst, 2012, p. 54), authorities in the space of justifications who form a "community of justification" (Forst, 2012, p. 19).[376] We show ourselves and other agents respect as justifying beings, as equal normative authorities who owe each other good reasons for how they treat one another and for the normative order to which they are subject.

Dignity is already the cornerstone at the level of normativity that precede political constructivism, that of moral constructivism which combines forms of factual and counterfactual justification, fostering discursivity among those affected and reflecting on justifiability within a comprehensive moral community of all human beings. Nonetheless, dignity is also central at the political constructivism, where human rights and basic justice principles ensure that those subject to a normative order are both governed by it and protected as equals: dignity, in this context, is connected to the idea that each person has an equal right to participate in shaping the norms, rules, and institutions that govern democratic society. Such participation is essential for a political order to be genuinely just, as it should enable the expression and respect of human dignity in practice (Forst, 2012, p. 139).

[375] "Failing to recognize someone's dignity by refusing to accord her second-personal respect may injure her in some way or other, but it cannot injure her dignity, at least not directly." (Darwall, 2013b, p. 16).

[376] "This context of practical justification is distinguished by its requiring reasons for actions, or for action-legitimating norms, adherence to which every moral person can demand from every other, even when those affected share no more closely identifiable ethical or political context. The justifying rea sons must be as concrete as the respective situation of justification is; here, they must be those that would be reasonably acceptable to persons in general. The connection between reason and morality emerges here: justifying reasons must in principle be accessible and agreeable to every reasonable person. In other words, a moral person must be able to take responsibility for his or her actions before affected others and also generally. The "community of justification" in moral matters is the community of all human beings as moral persons, and those concretely affected are, as representatives of this community so to speak, the primary addressees of justification. This does not mean that they are reduced to "generalized" others with no identity, but that they have, in all their particularity, the authority of the moral community of all persons "behind them" (metaphorically speaking)." (Forst, 2012, p. 19)

Forst thus subscribes to a notion of dignity rooted in the autonomous capacity of justification of human beings, albeit in his own distinctive manner. He agrees with Griffin's view that a notion of human dignity and normative agency stands at the center of human rights discourse but disagrees as to how to reconstruct this notion (Forst, 2010, p. 721).[377] Forst believes that the history of political struggles teaches us a particular political-moral lesson, presenting a political notion of dignity as a relational concept which refers to the social and political standing of human beings as agents of justification who are equal to one another. Forst thus enrich Griffin's concept of personhood by emphasizing an intersubjective dimension, proposing that humanity is inherently connected to the concept of status-respect through their relational existence. Correspondingly, the core question is not the foundation of human rights but their justification (Forst, 2010, p. 734)

This relational conception of dignity also seeks to address what Forst identifies as a significant misinterpretation within the Kantian tradition. He contends that "Kant overlooked this constitutive relation to others as a "fact of practical reason". He traced moral respect for others back to the wrong ground: the relation to oneself. But morality is in the first instance concerned with the dignity of other persons." (Forst, 2012, p. 55). Yet, many Kantian constructivists argue that this relational conception of respect is already inherent in Kant's work, precisely as a "fact of practical reason" (Bagnoli, 2007; Sensen, 2011).[378] Consequently, their understanding of dignity—and, by extension, Forst's—should be seen not as a re-semantization but rather as a recovery of the classical Kantian conception.

[377] "I agree with the claim that the "Enlightenment notion" of human rights is in need of an explicit philosophical justification, yet I think that the fact that the idea of natural rights was a polemical one directed against religious-political doctrines of the legitimacy of feudal social structures and absolute monarchy conveys a different and more determinate message than the one Griffin extracts from the historical account. As early as the sixteenth century (in the Dutch Revolution, e.g.), the discourse of human dignity was linked to the political question of opposing tyranny, that is, to protests against forms of political domination that did not regard those who are subject to such rule as persons to whom the exercise of political power had to be adequately justified. The political question of freedom and justice was essential to the formation of human rights discourse: the dignity in question referred to the status of persons who were no longer to be treated as servants, as second-class citizens, as not worthy of being regarded as normative agents to whom rulers owed reasons" (Forst, 2010, p. 721).

[378] "A failure to hold a person accountable for wrongdoing is tantamount to a failure to respect her dignity as a rational being. That is, the right to use moral sanctions (such as blame) springs from the recognition of the wrongdoer as having equal standing in the moral community, and therefore as somebody to be addressed as a wrongdoer, rather than obstructed as morally incompetent. On this view, then, blame of the wrongdoer, as a form of moral address, is a relation that presupposes the recognition of mutual accountability." (Bagnoli, 2007, p. 124).

Given that Kant's concept of dignity occupies a central place not only within constructivist discourse but also in the broader debates surrounding human rights, the next section will focus on examining and comparing two distinctive Kantian perspectives: Oliver Sensen's Kantian constructivism and Luigi Caranti's dignity-based approach.

3.2 SOME KANTIAN PERSPECTIVES ON DIGNITY

If human rights are described as "deriving from the inherent dignity of the human person", as outlined in the International Covenant on Civil and Political Rights (1966), then dignity seems conceptualized as an intrinsic value possessed by all individuals—a widely accepted interpretation in everyday discussions on the topic.[379] To provide a philosophical foundation and justification for this concept, many scholars turn to Immanuel Kant, who is frequently recognized as a seminal thinker who established dignity as a fundamental principle for respecting and valuing human beings, laying the groundwork for contemporary discussions on human rights and moral respect. If this philological interpretation holds true, scholars identifying as Kantian constructivists may face a dilemma: they must either abandon the notion of dignity as status or completely distance themselves from the Kantian label. I will illustrate that neither option is obligatory by presenting two distinct readings of Kant concerning the concept of dignity: the constructivist perspective, elucidated by Oliver Sensen (2011, 2017), and the realist interpretation, advanced by Luigi Caranti (2011, 2019a).[380]

By juxtaposing both readings, it will become evident what distinguishes the constructivist interpretation and how proponents can still maintain a Kantian concept of dignity: by rejecting the idea that dignity is an ontologically given status and instead seeing it as a status emerging from

[379] This perspective is so deeply ingrained that even dictionaries, such as the German Duden, define dignity (Würde) as a "value inherent in human beings that commands respect", and thus one should respect another person solely because they possess dignity (Sensen, 2011, p. 1).

[380] To be more precise, Sensen, although contributing to a collection explicitly titled "Constructivism in Ethics" (Bagnoli, 2013a), prefer to refers to his reading of Kant as "Kantian Constitutivism" (Sensen, 2017). However, as mentioned in the introduction to this project, constitutivism can effectively be seen as the most contemporary version of Kantian constructivism, from O'Neill and Korsgaard onwards. On the other hand, Caranti explicitly labels his interpretation as a realist reading of Kant but also specifies on various occasions that it is nonetheless a reading that does not align with realism 'all the way down' (Caranti, Forthcoming).

rational and public deliberation. Moreover, it will be shown how even a moderate realist reading of Kant can accept a non-foundational conception of dignity as a status recognized intersubjectively and not self-evidently, showing that accurate that accurate interpretations of Kant are not as distant from each other as they are often assumed to be. Ultimately, the primary objective of this chapter, in relation to our general project, is to show that Kantian constructivists do not need to abandon the concept of dignity as a status, nor must they relinquish their Kantian allegiance. Instead, they can refine their position by acknowledging the role of intersubjective respect in grounding dignity while preserving a coherent and philosophically robust constructivist framework.

3.2.1 A Constructivist Interpretation of Kant: Sensen

In the broader constructivist framework, Oliver Sensen's book, appropriately titled 'Kant on Human Dignity,' holds considerable significance for the specific question at hand. In this work, he delves into the intricate connection between respect for others and the concept of dignity, especially in light of the seeming absence of an idea of absolute value in Kant's oeuvre. This is particularly noteworthy as any external origin of value in relation to the moral law is deemed heteronymous.[381] The answer, in short, is that the faculty of reason itself prescribes the moral law, rather than discovering or reflecting an external moral standard. Dignity, instead of being an absolute value discovered as the foundation for respecting others, is the status attributed to human beings through our faculty of reason, recognizing them as autonomous agents deserving of respect. This marks a reversal of the traditional relationship, with value now dependent upon an a priori law of human reason, rather than the opposite: it is the Copernican Revolution of ethics.[382]

[381] Sensen emphasizes that according to Kant, if one does not commence with the moral law, then motivation would have to be derived from pleasure, but pleasure is relative and contingent, and therefore cannot produce the moral law. As we will see, Caranti subscribes to this view but asserts that this is precisely the reason why the moral law, inexorably linked to the autonomy of human beings, represents the absolute value that Sensen is attempting to eliminate.

[382] Sensen explicitly noted, following Engstrom, that Kant's views on the relationship between value and the moral law undergo a similar 'Copernican Revolution' in moral philosophy (Engstrom, 2009, p. 13): "The relevant feature is that Kant proposes a revolution in our way of thinking ["Revolution der Denkungsart"] (KrV B xviiif.). It is a natural way to think that human cognition must conform to the object. Similarly, it is a common way to think that moral laws must conform to something that has value or is precious. Kant reverses both relationships. In his theoretical philosophy Kant argues that the object must conform to a priori elements of human cognition. In his moral philosophy Kant argues that absolute value is dependent upon an a priori law of human reason" (Sensen, 2011, p. 26).

Sensen constructivist reading is text-based, it starts from the quantitative fact that Kant makes only a scattered use of the term "dignity" (111 times in all his texts)[383] and by noting that he never relies on value or dignity whenever he says that he justifies moral requirements. Kant regarded statements about value as prescription of reason: goodness is determined by what reason considers necessary. Therefore, something is deemed absolutely good or possesses intrinsic value only if reason deems it necessary without any conditions, and that in Kant is only a good will. What constitutes the worth of a human being is that he acts for the sake of the Categorical Imperative. Absolute inner value follows from the moral law, it is not the basis of it (Sensen, 2011, p. 51). Therefore, the true reason why we should respect others is because that is a direct command of reason.

But why should we follow a command of reason? Sensen attempts to reconstruct the entire Kantian framework by making a clear distinction between the categorical imperative and the moral law. The categorical imperative serves as Kant's formulation for accessing, comprehending, and applying the moral law—an inherent a priori principle. It functions as a practical tool enabling us to discern our duties, determining the actions mandated or prohibited by the moral law.[384] Rather

[383]"Of these there are only eight times when 'dignity' appears in conjunction with 'worth' or 'value'. One cannot plausibly substitute 'value' for 'dignity' in all 111 passages, but one can explain the eight value passages in terms of an elevation" (Sensen, 2011, p. 144). Moreover, the fact that the majority of time the word is used with another meaning testify for Sensen that the traditional association of dignity with absolute inner worth is not something Kant subscribe to: "The equation of 'dignity' with a morally relevant 'absolute inner value' does not square nicely with Kant's talk about the 'dignity of mathematics' (cf. KrV 3:323.09), the 'dignity of a minister' (ZeF 8:344.06 – 08), or 'the dignity of a teacher' (cf. RGV 6:162.19). In these instances Kant is not talking about something that should be pursued unconditionally" (Sensen, 2011, p. 143).

[384] Sensen reconstruct the crucial practical role the Categorical Imperative has in Kant's philosophy for coordinating the plurality of finite yet connected agents, who cannot avoid each other, but possess different needs and opinions. Relativism, the belief that different individuals can adhere to fundamentally different moral truths (or that no universal truth exists), is deemed either too early or too late, to cite Bernard Williams (1985). It is too early for a group with moral laws in common that does not know about other groups with different morals. However, once one knows about other groups with different rules, relativism is too late, as people with differing views have to find a way of living together (Sensen, 2011, p. 204). The categorical imperative, therefore, is not solely a procedure for deriving concrete duties; it is an essential practical tool for the coordination and mutual understanding among a multitude of finite rational agents, aiming to resolve conflicts. However, accepting this interpretation of Kant's philosophy necessitates further systematic justification. The pivotal question that emerges is whether the Categorical Imperative is indeed categorical as Kant intended. He defines 'categorical' as an imperative that holds true irrespective of personal desires. In contrast, a hypothetical imperative is conditional, relying on specific desires or needs; while a hypothetical imperative is subjectively valid, depending on whether its specific condition is fulfilled, it lacks objective and universal validity. The reading of the Categorical imperative as a practical tool may thus raise the objection that in doing so it will base it on the desire to avoid strife and isolation, rendering the imperative merely hypothetical and not categorical. If we follow the practical reading this objection is misplaced, since Kant seems to suggest that the option of living in isolation isn't a viable one, implying an inherent human inclination towards societal interaction, which could provide a basis for the categorical nature of the imperative. In this sense, the imperative can still be considered categorical

than being derived from a normative fact, such as a value, the categorical imperative stems from freedom in a descriptive sense, shaping decision-making and action (Sensen, 2011, p. 114). This is due to Kant's Newtonian conception of reality: since every form of causality in nature needs a law, the moral law is the descriptive law of freedom.[385] "[F]or Kant the imperative is a causal law. It describes actions that are governed by the causality of freedom. Kant does not invoke a value in order to explain why the Categorical Imperative is real and a valid command. Rather, the moral law is an in-built principle of a reason that has the property of freedom. The moral law is an operating principle of reason, so to speak. Reason automatically functions in accordance with this law. As the principle of non-contradiction guides one's thinking in theoretical matters, the moral law guides one's reason in moral deliberation" (Sensen, 2011, p. 115).

Sensen scrutinizes the diverse formulations of the categorical imperative, placing particular emphasis on Kant's reasoning for differentiating the Formula of Humanity from the broader Formula of Universality. He presents two distinct interpretations for why Kant introduces the Formula of Humanity alongside Universality: that of Klaus Reich and that of Henry Allison. Reich argued that Kant introduces different formulas of the Categorical Imperative to demonstrate its compatibility with, and to clarify the proper understanding of, the three Stoic ethical principles found in Cicero's *De Officiis*, popularized by Garve's translation and commentary. Instead, Allison has emphasized a more systematic reason, namely that rational agency involves acting on ends (Sensen, 2011, p. 111).[386] Sensen concurs with both readings, highlighting Kant's own acknowledgment of a desire to bring the imperative closer to intuition, influenced by Rousseau's ideas on honoring humanity. While the introduction of this formula remains speculative, Sensen's

---

because it applies universally in the context of human social existence, rather than depending on individual desires or conditions. We will return to this specific aspect of universality in section 5.2. See also Section 2.2.4 for a general constructivist interpretation of the categorical imperative.

[385] It must be noted that this emphasis on law was also prominent in the natural law traditions of thinkers like Grotius and Pufendorf. J.B. Schneewind, in his comprehensive analysis suggests that Kant's focus on law can be viewed as a response to these perspectives (Schneewind, 1998). For Kant, the foremost normative reality is indeed a law, but it is critical to note that this law does not originate from an external source – such as God, the state, or other individuals. Adherence to such an externally imposed command would constitute heteronomy. Instead, the law that Kant refers to is one that emanates directly from reason itself. This intrinsic law is self-governed, highlighting Kant's unique contribution to the understanding of moral and ethical principles.

[386] It is interesting to point out that Henry Allison (1986), and specifically his reciprocal thesis, according to which autonomy is not only a necessary condition for the moral law, but also sufficient (since they are reciprocal concepts), plays a central role in the argument developed by Caranti. However, Caranti employs this thesis to reach a conclusion seemingly opposite to that of Sensen—namely, that autonomy is precisely an intrinsic value.

interpretation suggests it reaffirms the fundamental command of universality, while encompassing a conception of agency and drawing this command closer to intuition (Sensen, 2011, p. 109).

In this context of the Formula of Humanity, for Sensen Kant does not deploy the phrase 'end in itself' to mean 'valued for its own sake', but rather to convey that humanity is an end because it is not merely a means in the clockwork of nature: it is free. According to Kant, humanity is necessarily an end for everyone because it is free from all laws of nature. An end in itself, in this sense, is an independently existing end, free regardless of anyone's will (Sensen, 2011, p. 110).[387] The Formula of Humanity expresses that what is an end in itself should also be treated as such, namely that free beings ought to be treated as such. Kant justifies this imperative by asserting that it is directly commanded by the Categorical Imperative—a straightforward directive of reason. However, the underlying justification for this command warrants examination (Sensen, 2011, p. 113). Sensen argues that the Formula of Humanity shares the same justification as the core of the Categorical Imperative, that one's maxim should be able to be willed as a universal law. This also means, according to Sensen, that one can use the Formula of Humanity (instead of the stricter method of the Categorical Imperative) to derive concrete duties.

Sensen, thus, interprets Kant's usage of the term 'humanity' as a reference to freedom or the capacity to be morally good, distinguishing two forms of humanity, namely "homo phenomenon" and "homo noumenon" (Sensen, 2011, pp. 128–129). The former represents a human being in the state of sensibility, while the latter represents an idealized version of a human being as they ought to be according to reason. The distinction is not metaphysical but concerns how humanity is experienced, and the moral idea one can have of oneself. The interpretation suggests that Kant's concept of humanity in the Formula of Humanity aligns with the capacity for freedom under the moral law. This is a crucial distinction, since freedom under the moral law doesn't necessarily imply following it, so we must be clear if what we are respecting is the morally good will or the capacity for it. In the Doctrine of Virtue, Kant emphasizes that one should respect the self-esteem of others, grounded in their freedom: this self-esteem is considered necessary, stemming from the

[387] "Human beings really are free and ends in themselves, according to Kant. One difference therefore is that I do not read 'end in itself' as 'valued for its own sake', but as 'not merely a means to the will of another'." (Sensen, 2011, p. 110).

duty to esteem oneself. Kant argues against servility, asserting that every individual can measure themselves equally in terms of moral worth.

The rationale for respecting others, thus, lies in the fact that every individual can strive for moral worth, achievable through adherence to the moral law (Sensen, 2011, p. 131).[388] Kant links the duty of self-esteem to the moral law and emphasizes that one should not dispose of oneself, recognizing an obligation to the noumenon or freedom. The duty to respect others' freedom, grounded in their capacity for morality, is derived from the Categorical Imperative, ensuring that maxims can be universal laws[389]. Sensen has characterized this interpretation of the obligation to respect others as a "thin reading" of the Formula of Humanity, that stands in contrast to a "thicker reading", which posits that the formula is based on a substantive value claim (Sensen, 2011, p. 135); the rationale behind respecting others, according to this interpretation, lies in its alignment with the command of the Categorical Imperative—a direct dictate of reason. When evaluating whether one's maxim could become a universal law, the consideration extends to its adoptability by all others, embodying the universal principle. The thin reading ensures coherence in Kant's texts, aligning the Formula of Humanity with the Categorical Imperative. Moreover, it harmonizes

[388] "To say that one should respect the homo noumenon is just another way of saying that one should respect freedom. […] But this is not the same as saying that what one should respect in others is a morally good will. The homo noumenon is freedom or the capacity for morality. […] What one should respect in others is therefore their freedom and their striving for morality" (Sensen, 2011, p. 131). This is not a circular or an exemplarist justification of why we should follow reason—that is, it is not "we should follow reason because the *homo noumenon* acts morally"; the focus is rather on the capacity for freedom inherent in rational agents and our capacity to recognizing its internal genuineness. We will return to this notion of genuineness on the next page, but it is useful here to clarify that, in Kantian terms, freedom is the broader capacity to act independently of inclinations or external determinants—the noumenal self that can determine its actions through reason (for example, refusing dessert for health reasons even though one's appetite desires it). Autonomy, by contrast, is the capacity to exercise freedom for moral matters according to the moral law, namely through the self-legislation of universalizable principles. When Sensen emphasizes respect for the *homo noumenon*, he is not claiming that others are automatically morally good; rather, he stresses that their freedom—their capacity to choose and (potentially) act morally—is what commands respect. These points will also be revisited in the next section when discussing Kant's distinction between humanity and personality.

[389] This perspective prompts questions about why the imperative doesn't extend to non-rational beings, raising considerations about arbitrariness and speciesism in Kant's ethical framework. Sensen address this question directly: "While animals are not to be tortured, according to Kant, one might still think that Kant's view does not protect them properly: for their own sakes. But this is not a charge that is particular to Kant's views on animals. Even duties towards other human beings rest on a duty to oneself, not on a factum (e. g., a value) the other possesses [...] Rather the duty to respect others rests on the Categorical Imperative. Since it is the imperative that commands respect for others, one first has to be constrained to follow the law springing from one's own reason" (Sensen, 2011, p. 134). Whether some animals, like chimpanzees, elephants and dolphins might have freedom or reason, that doesn't change how one should treat them according to Kant: we should respect others independently of who exactly these others are, and this is a duty to ourself (Sensen, 2011, p. 135).

well with Kant's broader philosophy, emphasizing a priori principles as seen in the first Critique (Sensen, 2011, p. 140)[390]. In asking whether one's maxim could be a universal law, one also ask whether the maxim could be adopted by all other and this is what threating others never merely as a means amount to.

To summarize Sensen's argument up to this point, which corresponds roughly to the part I (chapters 1 to 3) of his book: Kant's Categorical Imperative is a practical application of the moral law, the core of which is universalizability (that we should only act according to a maxim that we can will to become a universal law). There is no absolute value prior to the moral law since Kant considers statement about value to be nothing more (and nothing less) than prescriptions of Reason: good is what Reason deems necessary, and therefore only something that Reason deems necessary without condition has inner value. The further question is then: why should we follow the command of reason? The answer, according to Sensen's interpretation, lies not in some external incentive or prior value, but in the very structure of rational agency and the genuineness of the moral law (Sensen, 2011, p. 117);[391] the law is "genuine" because it is not imposed from without, nor conditioned by subjective inclination, but arises from our own rational will—our "proper self" (Sensen, 2011, p. 23).[392] Precisely because it has this genuine, internal source, we recognize it as valid, not a heteronomous demand, and take an interest in it, not vice-versa.

[390] While some may desire a more depth moral theory, the thin reading aligns with Kant's emphasis on freedom as the fundamental and essential element of morality. The requirement to treat free beings as free stands out as a modest yet beneficial stipulation. Instead of anticipating a universal solution applicable to all societies, Kant's framework accommodates diverse interpretations of the good life, highlighting the coexistence of different traditions and the respect for humanity in its varied expressions. This adaptability is seen as a strength in Sensen's perspective, in contrast with Hegel's viewpoint: "The traditional charge that the imperative is empty (cf. again Hegel 1820, §135) is true to an extent, but this is a strength rather than a weakness. Why would one expect a person in 18th century Prussia to come up with one sentence that will decide once and for all every moral issue (including the ones about technologies that have not yet been invented)? Instead Kant gives us a framework within which the details are to be worked out by each group" (Sensen, 2011, p. 206).

[391] "[I]n realizing that the moral law is genuine (i.e., not conditioned by a subjective inclination), one is interested in it, according to Kant. Is Kant too optimistic on this point? Again my aim is not to defend him systematically. Kant himself does not seem to argue for the claim that even the most hardened scoundrel wishes to be moral" (Sensen, 2011, p. 117).

[392] "Kant talks about the "proper self " in the context of explaining why one should take an interest in the moral law […] The law is valid because it has a genuine source that is not external; an external source could not ground a necessary and universal law (see Sections 1 and 2). Since the law is genuine, one should take an interest in it. There is no further argument that an intelligible world has value." (Sensen, 2011, p. 23).

In this way, we acknowledge that the *homo noumenon*—the autonomous self capable of acting according to reason—is good insofar as it acts from freedom and in conformity with the moral law. For Kant, the good is dependent upon the right, and it is not humanity that has inner worth, but morality. There is therefore no independent worth that could ground the requirements to respect others besides the moral law and its practical application of the categorical imperative. One should respect others because by reasoning we arrive to the maxim of not exacting oneself above others.

If Kant doesn't hinge the foundation for respecting others on a conception of value and instead ground it on a direct command of reason, how does he integrate the concept of 'dignity'? This is the central question in the context of the human rights debate we are discussing. The short answer, according to Sensen's interpretation, is that Kant intertwines the concept of dignity with duties towards oneself.[393] According to Kant, dignity is not an inherent feature (such as an intrinsic value) that autonomously creates rights. Rather, the rights one can claim follow from a duty of the agent: "the concept of claim rights is preceded by the concept of duty. Someone can claim rights by reminding the agent of his or her duty to follow the Categorical Imperative. Accordingly, the 'innate right of each' is a right to freedom that can coexist with the freedom of everyone else in accordance with a universal law" (Sensen, 2011, p. 169).[394]

The term dignity is thus employed by Kant to emphasize the superiority of morality over other forms of behavior. However, this term alone does not establish any specific moral obligation: respect for the moral law is the proper moral motive. Therefore, dignity occupies a secondary role in Kant's ethics, compared to its prominence in the contemporary understanding of the concept.[395] Nevertheless, this does not render either dignity or Kant's ethics irrelevant; it simply underscores

---

[393] "In sum: In accordance with the traditional paradigm of dignity, Kant uses 'dignity' predominantly in the context of duties towards oneself. As respect is connected to the duty to oneself to follow the moral law, dignity appears foremost in connection with duties towards self"(Sensen, 2011, p. 172).

[394] This way of proceeding is not, for Sensen, a "new justification (in the sense that the noumenon has a higher value than the phenomenon). Rather a human being owes respect (and feels respect) for the noumenon. […] So, if one should respect others – as is required by the formula of humanity – then one should respect their freedom in virtue of which they have a duty of self-esteem (as expressed by freedom's moral law)"(Sensen, 2011, p. 132).

[395] "This explains why Kant uses the term merely sporadically, and not at all in places one would expect him to if he had the contemporary paradigm in mind. As such dignity is not a concept that carries any justificatory weight. Kant uses it sometimes to express his core ideas. But this should not be surprising. Kant often takes popular concepts of his time and subordinates them to his theory (e. g., the enthusiasts of his time or religious terms like holiness, cf. KpV 5:86 f., or Stoic conceptions like dignity)" (Sensen, 2011, p. 173)

that dignity does not hold a foundational position. Instead, it suggests that morality takes precedence over actions motivated by inclinations. Understanding Kant's use of dignity in this context lends coherence to his texts and clarifies why he references 'dignity' infrequently, and notably, omits it in passages where one might anticipate its presence. In this sense, dignity is not as a self-evident value-concept (as moral realists would hold) but the result of rational moral deliberation—an idea constructed within the framework of moral reasoning itself. "Dignity is explicitly said to follow from the demand of the Formula of Humanity, it is not the ground of it." (Sensen, 2011, p. 202).

In short, according to Sensen reading, the place of human dignity in Kant's framework "is more modest than it is often taken to be; but it is not insignificant. Human dignity does not ground moral imperative, but human beings have dignity because they are free and so bound by moral imperatives" (Sensen, 2011, p. 213). Sensen reverses the relationship between dignity and respect. It is not that someone should be respected because he has dignity, but he has dignity because he should be respected: a human being has dignity because he should be treated as an end in itself, not the other way around. Sensen's interpretation explains how one can read Kant's view on dignity, respect and humanity as end in itself without reference to a value as a foundation.[396]

3.2.2 A Realist Approach to Kantian Dignity: Caranti

In contrast to Sensen's constructivist interpretation, Luigi Caranti develops a realist reading of Kant across the three most crucial domains of his philosophy: epistemology, morality, and politics (Caranti, forthcoming).[397] Comprehending Kant's realist approach to his theory of

[396] "[R]espect for others is a task that is required of us. Kant's conception of dignity is important because it emphasizes our sense that this task is nobler and more sublime than the selfish pursuit of our own desires" (Sensen, 2011, p. 212). This nobility (or dignity) of the task of respecting other does not possess intrinsic value in itself, independent of reason and the moral law. We are the source of value, not as end in themselves, but for the constitutive principles of our reason.

[397] The paper I am referring to, provocatively titled *Kant, the realist*, was presented during Rainer's Forst Kolloqium the 4th of February 2024 and will be published in some revised form in the near future. Despite the catchy title, Caranti acknowledges that Kant doesn't neatly align with either realism or anti-realism "all the way down", similarly to the constructivist reading. Nonetheless, his interpretation diverges from constructivists by asserting that Kant maintains the belief in the existence of moral value in the world, at least for a specific entity: the autonomy of a rational agent. Autonomy, defined as the capacity for moral action, not only functions as a necessary condition for the moral law (which, from this standpoint, is the final step in establishing value) but is also deemed sufficient. In line with Henry Allison's reciprocity thesis, autonomy is considered an absolute value in this sense (Allison, 1986).

knowledge, as opposed to the dualistic perspective articulated by Paul Guyer, provides for example an alternative to the prevailing naturalist paradigm.[398] This has direct implications for Kant's moral philosophy, as it becomes challenging to view ourselves as truly autonomous in a strictly naturalist world. We consider ourselves existing as entities within a realm governed by the principle of causation, while simultaneously acting freely in the strong incompatibilist sense essential for grounding morality. But what is the rationale to believe that we are not just potentially free but genuinely free? Caranti underlines how for Kant, this justification stems directly from a "fact of reason" – an immediate awareness of the authority of the moral law. It is the consciousness that duty is an imperative for us because it originates from our own reason, not from some dispensable external source. The term 'fact of reason' is metaphorical, as it doesn't denote something encountered in the sensory world. It signifies that reason is not in contact with the external world but with itself and the supreme law governing its practical use.

While some constructivist positions like that of Paul Formosa (2013) or Fred Rauscher (2015) are all very plausible both as an approximation to what Kant meant and as philosophical theories in their own right, Caranti's impression is "that a crucial realist element is still missing" (Caranti, forthcoming, p.5).[399] Caranti directs attention to the renowned 'remark' in the second critique, where Kant tackle the question of how is consciousness of the moral law possible by providing two accounts of how we immediately grasp it, emphasizing an awareness of its validity for us. Initially, Kant suggests that this awareness mirrors the way we intuitively grasp the validity of theoretical principles, such as causality or specific mathematical theorems, like the Pythagorean theorem. In this sense, the truth that one cannot borrow money without the means to repay it

[398] "[A]ppearances and things in themselves should not be considered, respectively, as mental entities and 'real things' existing outside of our mind'. This would make Kant a horrific combination of Berkeley and Descartes, with the famous synthesis of the rationalist and empiricist school brought about by the critical philosophy consisting in combining phenomenalism and a dogmatic affirmation of the existence of a set of things that 'must be there' even if we will never know them. […] If one opts for Kant's realism, properly understood, then it is simply not true that in principle all questions about the world can be answered by science, not even at the end of time, our grasp of reality will never be freed (and does not need to be) from the mark of subjectivity highlighted by the Copernican revolution. If the alternative reading of Kant is adopted, however, Kant would have very little to say to contemporary science, given that the central message of his epistemology is that all we can know are mental entities." (Caranti, 2024, 2).

[399] "And my suspicion is that it is missing because unconsciously these authors have not digested the whole impact of the Copernican revolution or perhaps equally unconsciously they still operate with a two world view of the transcendental distinction" (Caranti, forthcoming, p.5).

becomes as clear and distinct to Kant as mathematical truths, and perhaps even more immediately apparent. [400]

Yet, the most crucial Kant's parallel is that with freedom (KpV, 5:30), when he asks us to consider whether, in the face of a life-threatening situation—such as a prince coercing us to give false testimony against an innocent man—we would have the moral strength to resist intimidation and face the consequences. According to Kant, none of us can know with certainty whether we would truly resist in such a scenario. However, it is abundantly clear that each of us understands a) what should be done and b) that we could carry it out, irrespective of the repercussions. In simpler terms, when confronted with a moral dilemma of this nature, everyone's response will invariably combine certainty about what should be done, and uncertainty, perhaps skepticism, regarding our courage to do what is right, but still undiminished certainty that we are capable of doing it.

The interaction between our consciousness and the imperatives of morality, the Kantian "fact of reason", 'proves' that our freedom or capacity for self-determination extends beyond merely responding intelligently to needs or desires. In the negative sense, we can act even when not driven by any empirical desire. Positively, we can discover a pure interest in the righteousness of the action that morality demands. These indeed represent the two facets encapsulated in the concept of autonomous agency: independence from external factors, including our sensible drives, and yet the ability to find a motive to act 'within ourselves'—in our own reason— potent enough to propel us, surpassing any other incentive. This brings us to Kant's distinction between *ratio cognoscendi* and *ratio essendi*: the moral law serves as the ratio cognoscendi of freedom (our only access to freedom is through the consciousness of duty), while freedom acts as the ratio essendi of the moral law (if freedom were not real, there would be no moral law, no moral value— absolute or otherwise— in the world).

[400] This seems to mark a return to the Rational intuitionism that Rawls and the entire Kantian constructivist reading sought to distance themselves from. Caranti acknowledges this but emphasizes that the immediate consciousness of the moral law does not entail intuiting an object in the world, analogous to Plato's notion of intuitions regarding mathematical objects. Instead, our immediate intuition pertains solely to recognizing our condition of being constrained by the unquestionable validity of a norm. Furthermore, the awareness of this constraint does not merely signify something confined within the realm of our rationality, nor does it render the world entirely indifferent; rather, it serves as an indication of our objective nature (Caranti, forthcoming, 11). Constructivist can equally endorse all these points and—anticipating the conclusion of this section—this suggests that the divergence between Sensen and Caranti is not as deep as it has been supposed.

Therefore, in contrast to certain anti-realist and constructivist interpretations that ground morality in the nature of reason—internal to agents and valid only subjectively—Caranti, moving away from a two-world view, argues that, although not empirically cognizable, freedom is nonetheless recognized as a defining characteristic of humans within the known world. He maintains that this freedom is not validated by theoretical reason but by practical reason, affirming its integral role in our moral framework. To elucidate this point, he reconstructs Kant framework by distinguishing two distinct conceptions of freedom: practical freedom and autonomy. The entirety of human behavior is subject to free rational deliberation: this implies that all voluntary actions originate from an individual's free assessment of a specific maxim, conceptualized as a subjective rule of action. Consequently, humans employ their reason to scrutinize the moral or prudential viability of a given maxim and are free to embrace or reject it. Kant articulates this fundamental idea by invoking the concept of "arbitrium liberum", distinguishing it from the "arbitrium brutum" typical of animals (KrV A533–534/B561–562).

Humans autonomously establish the purposes of their lives, not being compelled or entirely determined by desires and needs. Desires and needs, at most, suggest a certain course of action; nevertheless, a free decision by the individual to endorse or reject such a path, thereby adopting it as one's own end, is always indispensable. Kant occasionally refers to this capacity of humans to posit ends that are not imposed on them by external forces or internal passions as "freedom in the practical sense".[401] However, "for the mature Kant, an agent whose freedom is limited to this ability is free but irremediably heteronomous. The agent is free because her inclinations (no matter how strong) do not exhaust the causal story behind her actions (it always takes a free rational act of endorsement), yet the agent is heteronomous because inclinations are necessary components of the motivational story behind the action" (Caranti, 2011, p. 38). This is why Kant distinguishes an

[401] Caranti holds that Kantian notion of practical freedom comes very close to what one nowadays would call rational agency, and indeed Kant would consider humans as rational agents precisely because they select the rule that guides their behavior (they do not act randomly), and they are free to select such a rule. To give an obvious example: if I am thirsty, and I see no reasons why I should resist or delay the satisfaction arising from the extinguishing of my thirst, I can freely select the maxim "any time an agent X is thirsty, she should drink." This is indeed particularly demanding, since according to Kant it is only on this condition that I can say that I perform any truly free action, as opposed to "something in me led me to act" or "that particular objective state of affairs evolved in a certain direction." Any time we are not deliberating on the maxim to follow, we are not freely acting in this practical sense. In this reading, to act autonomously adds an extra layer of difficulty, by removing any possible empirical influences on the selection of the maxim. On the relationship between empirical inputs and autonomous reasoning in a Kantian constructivist sense, see section 2.2.4.

even nobler sense of freedom, namely autonomy. Given that any form of agency presupposes that the agent is motivated by some interest, it implies that moral or autonomous agency is driven by the sole form of pure, non-empirical interest accessible to humans—respect for the moral law. According to this interpretation, autonomous behavior is synonymous with moral behavior.[402]

Moreover, Caranti underline how practical freedom and autonomy are what Kant is pointing at when distinguishing between humanity [Menschenheit] and personality [Persönlichkeit]: "Kant had in mind that it is our capacity for moral behaviour – personality – that establishes humans' status as ends in themselves, thereby entitling and binding them to morality. On this interpretation, even when Kant suggests that it is humanity that is to be seen as an end in itself, what he really means is that it is humanity insofar as it also embodies the capacity for morality, that is, personality." (Caranti, 2023, p. 3).[403] Thus, practical freedom showcases our rational nature, elevating us above the natural world, but it is autonomy, seen as a necessary attribute for a will to become a good will, that establishes our superior worth as person compared to other entities in the sensible world. In this sense, this autonomy is intrinsically valuable, not per se, but because it reveals that we all are at least capable of something extraordinary, truly moral behavior, that – we assume – exacts respect. We are not merely self-masters, but also, and most significantly, potentially righteous ones. We are not simply free; we are free to choose a path of integrity and mutual respect.

Caranti has further refined and affirmed his personal theory on the foundation of human rights, labelled precisely the "dignity approach", by building on the Kantian reading we have just discussed (Caranti, 2019a). However, he also intentionally departed from certain elements of the traditional philological interpretation of Kant, challenging for example Kant's assertion that "the capacity for moral agency is a peculiarity of humans",[404] and, on the other hand by holding that "humans act autonomously not only when they follow the Categorical Imperative, but also when

[402] Although, obviously, moral behavior is only a subset of a practical free behavior which encompasses both autonomous and heteronomous actions (KpV 5:33-34).

[403] Caranti is arguing, against Korsgaard and Wood, that "personality, not humanity, is the ground that Kant perceives as the basis of humans' (and rational beings') status as ends in themselves, and thus that personality, not humanity, is the source of moral normativity. It also argues that personality, at least on Kant's mature theory, is a capacity that is not already contained in humanity or inferable from it" (Caranti, 2023, p. 3).

[404] On how constructivists address this question, we will return in section 3.4.

they adopt different moral formulas – like the Golden Rule – or act virtuously in an Aristotelian sense." (Caranti, 2018, p. 111).[405] As he puts it: "we want to sever Kant's account of autonomy and dignity from his analysis of morality as necessarily grounded on the Categorical Imperative. Although at some cost, this will give our approach the latitude necessary to consider actions performed under a variety of moral rules as instances of authentic moral behaviour. This is crucial in constructing a notion of human dignity that can be accepted outside the small circle of Kantians" (Caranti, 2019a, p. 58).

Besides this explicit departure to develop his own theory, Caranti's main argument in his reading of Kant can be succinctly summarized as follows: the concept of dignity in Kant does not lie merely with humanity, the capacity to set an end, but with personality, the capacity for moral action. Nevertheless, dignity is not dependent on performance (criminals don't lose their dignity), since it is not only a necessary condition for the moral law but also sufficient, and in this sense, it is an absolute value (Henry Allison's reciprocity thesis). However, this absoluteness does not necessarily imply an unconditionally foundational nature,[406] rather it can be understood in a relational manner, as Caranti has endeavored to elucidate in response to his critics.[407] This relational aspect emerges from the intertwining of autonomy and respect. Dignity, as the embodiment of moral autonomy, grants individuals the right to be respected, which entails a guarantee that they will not be treated merely as means to an end. And precisely because we have

[405] Caranti maintains, against the notable interpretations of Hermann (1993) and O'Neill (1989), that the moral law, as expressed in the Groundwork, was for Kant a more philosophically rigorous version of the Golden Rule, which people can follow even without knowing the categorical imperative. He thinks that Kant would not claim that individuals who adhere to the Golden Rule are not autonomous due to their lack of exposure to his moral philosophy. The aim of the Groundwork, Caranti suggests, is to provide "a better formulation of the moral law, not to explain to all the decent people in the world that they fooled themselves to be truly moral simply because they had not read that philosophical work yet." (Caranti, 2018, p. 111). Still, Caranti's attempt to detach autonomy from the moral law is more radical, in the sense that whoever consciously follow a tradition or a religion in orienting their lives, even if characterized by some controversial moral principles, may be autonomous (Caranti, 2019a, p. 58).

[406] "The general idea of my approach is that dignity is both an intrinsic and a very important value. So important that it trumps other moral values (happiness, security, utility). It is thus absolute in the sense that it is freed (ab-solutus) from the competition with these other values. But certainly it is not absolute or unconditional in the sense that one cannot weigh conflicting claims to the protection of the very same moral value."(Caranti, 2018, p. 114).

[407] "In the context of my argument I do make autonomy as a relational property. I put it in terms of a difference of degree rather than in kind between human and animal autonomy. Hence, I guess that I am committed to say that dignity is a relational property (probably this is not the case with Kant, as Tonetto argues against Sensen). But again, I do not see what major problems this has for my ability to secure humans a special treatment compared to other species, or at least to move the first step in that direction."(Caranti, 2018, p. 113).

this capacity, precisely because morality is within our reach, we are entitled to an amount of respect unfettered by contingent circumstances: in this sense this capacity is something intrinsically valuable that serves as the foundation for respect.[408]

3.2.3 Towards a Kantian Convergence on a Relational Status

We have shown how Kant concept of dignity can be interpreted very differently depending on what reading of Kant are we subscribing to. Both Sensen and Caranti agree that Kant is neither a realist nor an anti-realist in the strict sense, but they disagree on what is his intermediate position: while Sensen argues that Kant is a constructivist, Caranti still define Kant as a moderate realist. However, given the potential drawback of this distinction being perceived merely as a terminological dispute without substantial underlying content, it is beneficial to delve into the key points and passages of the Kantian schema leading to the imperative to respect others, to check what are the substantive differences between the two proposals.

Any theory of dignity and respect claiming a Kantian foundation must originate from the command that emanates from reason. The moral law articulates itself clearly and uniformly for everyone; it possesses an inescapable authority, constituting a fundamental element of constitutivism acknowledged by both Sensen and Caranti. Beyond the inescapable nature of the moral law, we continually retain the capacity to make exceptions for ourselves—a facet of human nature that, according to Kant, is indelible, though not incurable. The autonomous reflection on the moral law does not engage us in a solipsistic procedure detached from the real world; on the contrary, interactions with others is necessary to comprehend the existence of the moral law itself, an intersubjective dimension inherent in Kant's concept of reason that both Sensen and Caranti recognize.

Precisely because human beings are free and so bound by moral imperatives, they are end in themselves and have equal dignity. The key divergence between the two perspectives lies in the

[408] "Autonomy is not to be understood merely as the ability to choose one's path in life, or as the ability to be rational in the sense of purposive agents. With Kant, we refer to a capacity distinct from and 'higher' than practical freedom. We have in mind the ability to act under self-imposed moral constraints. This capacity […] shows us as worthy creatures, and reveals the deepest and most stable layer of human value. Reflecting on our autonomy, we turn out to be beings with this fascinating feature: being able to silence all natural impulses, even the strongest instincts of survival, and act from our conception of duty" (Caranti, 2019a, p. 57).

value attributed to dignity. Caranti argues that dignity is intrinsically valuable, while Sensen firmly maintains that Kant does not endorse a conception of absolute value that is antecedent or independent from the moral law. Nevertheless, Sensen explicitly says that there is no disagreement between his reading and the proposals of scholars like Christine Korsgaard, Allen Wood, Paul Guyer, Richard Dean and Samuel Kerstein.[409] "On these readings value is not necessarily a distinct metaphysical property. There is only a disagreement if one understands this value as being priori to or independent of the moral law, or if one ground the requirement to respect other others on the value of a good will, not on the moral law (Sensen, 2011, p. 52).[410]

There is a subtle yet significant difference in delineating dignity as a metaphysical value-concept and asserting that our capacity as moral agents is something a perfectly rational being would value. However, even this type of proposal, according to Sensen, falls short in deriving a moral conclusion (such as the obligation to respect others) from non-moral premises, like the capacity to set ends in general, as undertaken by Korsgaard, or from the morally good will, as argued by Dean or Kerstein. The main problem remain unchanged: the claim that we possess absolute value because we are the foundation of value doesn't logically follow, as Sensen has contended, providing various examples (such as the analogy that if something is only perceived as funny by humans doesn't imply that humans are inherently funny).[411] Despite our inclination to value our autonomy and capacity for a good will, Sensen would still say that we are not respecting others for this capacity, since "duties towards other human beings rest on a duty to oneself, not on a factum (e. g., a value) the other possesses [...] Rather the duty to respect others rests on the

[409] On the conception of values used by constructivist, see section 2.3 of the previous chapter.

[410] Sensen categorizes these authors into two groups. The first group argues that one should respect others due to the value of their pre-moral capacities (Korsgaard, Wood, Guyer), while the second contends that we should respect others for their morally good will (Dean, Kristensen). Given our discussion thus far, one might assume that Sensen would place Caranti in the same group as Dean and Kristensen. However, I believe this doesn't do justice to Caranti's distinction between practical freedom and autonomy, and how the potentiality of being moral constitutes a status distinct from contingently possessing a good will. This is the main reason why, as I will contend in the next page, Caranti's and Sensen's proposals can find a common ground.

[411] "If a thing only has value if it is valued by human beings, then it does not follow that human beings must have an absolute value, as has often been noted. In general, it does not seem to be the case that if one thing confers a property onto another, the first thing has this same property unconditionally. For instance, a university president can confer a PhD title onto a student, but this does not mean that the president has a PhD or an unconditional PhD. […] Here I am assuming that a line of words is not funny if it were the only thing that exists in the world. Rather it is only funny if it is funny to an observer. But if something is only funny if it is funny to human beings, then it does not follow that therefore human beings are funny in themselves." (Sensen, 2011, p. 64).

Categorical Imperative. Since it is the imperative that commands respect for others, one first has to be constrained to follow the law springing from one's own reason" (Sensen, 2011, p. 134).

A will is good if it follows the categorical imperative for its own sake, but the good will of others is not the reason why one should respect them: for Kant even a vicious person deserves respect as a human being, because we should respect all human beings equally, independently of their possession of a good will (Sensen, 2011, p. 92). Caranti however put the focus not on the possession of the good will, but on the possibility that every human beings has to be reach it, namely autonomy, so it does not fall into the critique that Sensen moves for example to Dean[412]. This explain what Caranti intends when it says that dignity is an intrinsic value: being autonomous agent is the foundation of the authority (for us and other rational beings) of the moral law (ratio essendi) while the authority of the moral law (which we 'perceive' through the fact of reason) makes us aware of our autonomy (ratio cognoscendi). In this respect, Caranti evades Sensen's critique, as this understanding of dignity doesn't position it outside the moral law, but place it right in the middle of nexus shaped by the interplay of freedom and reason.

The very fact that both Caranti and Sensen underlined the distinction between a morally good will and the capacity of it, opens the possibility of a common ground that both authors could potentially agree upon: Caranti might concede that dignity can be categorized as a status-concept, as proposed by Sensen's interpretation, but only if we perceive it as a status based on the value concept of autonomy, as advocated by Caranti. Despite both authors claiming to interpret Kant in seemingly opposing ways (the constructivist and the realist reading) they can arrive at this shared understanding that significantly deviates from the standard discourse in the human rights debate: dignity is not a distinct metaphysical property but rather the status attributed to every human being due to their potential ability to heed the command of reason and act morally.[413]

---

[412] Dean tries to argue that one should respect all human beings, and not just the one with a good will, by adducing for example that respecting vicious human being might be needed to give them a chance to better themselves or that not respecting them could have a corrupting influence on the agent's character. This is neither enough egalitarian nor Kantian for Sensen (Sensen, 2011, p. 92).

[413] On the subtle relation between realism and constructivism regarding this specific Kantian theme, see Formosa (2013). "For the law to be binding on us we don't have to do something. We simply have to be something, namely beings capable of acting autonomously on the basis of reasons, and we are bound by the categorical imperative because it is constitutive of practical rationality. This is a constructivist view and not a strong realist view because it is not based on an independent conception of value which precedes and grounds the law. Rather it is based on the constructivist claim that practical reason constitutes a conception of objective value, namely of the absolute worth of rational agents and the conditional worth of their rational ends" (Formosa, 2013, p. 196).

A final observation is warranted, while we summarize and reconnect to the main argument of the chapter. Both interpretations position the source of normativity of the Kantian concept of dignity within our capacity for morality, thus rooted in personality rather than humanity; we have elucidated why this doesn't imply conferring rights solely to morally virtuous agents, as the grounding of these rights lies in the potentiality of this characteristic, not its actualization. Nonetheless, this move seems to center the concept of dignity and the corresponding respect it demands primary on (practical) freedom (specifically, the autonomy inherent in being a moral agent) rather than on the concept of equality (which is only the first universal right derived from the innate right to freedom). This result remains susceptible to common non-Western-centric critiques that prioritize equality over freedom, as well as to various non-liberal approaches such as communitarianism, which place greater emphasis on communal values over individual liberty.[414] We will return to Kant's conception of the innate right to freedom in the following chapter, through an analysis of Arthur Ripstein, in order to assess whether Kant is best understood within a jusnaturalist or juspositivist framework.

3.3 ALTERNATIVE FOUNDATIONS

We have explored the possibility of conceptualizing dignity in a way that departs from the traditional value-based framework, allowing human rights to be grounded and justified in a non-foundational manner (3.1). Then, we demonstrated that this approach aligns with Kant's conception of dignity, which is often misinterpreted in discussions of dignity and human rights (3.2). Nonetheless, there are still scholars who opt for alternative concepts as the foundations of human rights, rather than dignity. To categorize these diverse proposals within broader frameworks, it is widely common to refer to a tripartition outlined by John Tasioulas in "On the

[414] Regarding specifically the communitarian critique, Rainer Forst has sought to demonstrate its misplacement and the creation of a false dichotomy between the rights of liberty and equality, which is circumvented by the centrality of the right to justification. "The rights of individual liberty - the right to respect for the person and for the personal freedom to act – are right to liberty that cannot be restricted by reciprocally and generally justified norms; in fact, those norms protect them. Liberty rights are thus not "innate" or primarily justified on the basis of "highest-ordered interests" (Rawls); they are moral rights of "reason" in the sense that they cannot be reasonably disputed inasmuch as persons recognize one another as authors and addresses of the norms to be justified. […] Therefore, what is foundational is an individual basic right to justification, the right of moral persons to veto, as it were, certain norms (or actions)." (Forst, 2002, p. 81).

foundations of Human Rights" (Tasioulas, 2015), where he distinguishes precisely between: 1) Reductive or instrumental foundations, 2) Orthodox or non-instrumental foundations, and 3) Political or practice-based foundations. Once again, this distinction suffers from the same issues of blurred boundaries previously discussed in categorizing different approaches to dignity.[415] While it can still serve as a useful orienting map, for the purposes of this project I will focus solely on the distinction between Reductivists and Orthodoxists.[416] This dichotomy is crucial for understanding the divide between instrumental (reductive) and non-instrumental (orthodox) justification approaches, and for determining the appropriate placement of a constructivist perspective, as already anticipated in paragraph 2.2.3.

3.3.1 The Reductive View: Human Capacity and Needs

Reductivists (or instrumentalists), conceive human rights as useful means to achieve certain aspects of human lives deemed inherently valuable and deserving of protection. They argue that understanding human rights is most effective without an essential reliance on the concept of a moral or natural right, but rather by focusing on some 'evident' characteristics or interests of human beings. In this sense, human rights are justified as means that permits to satisfy these characteristics. When it comes to specifying what these features are, Cruft and others have

[415] Tasioulas himself was categorized within the instrumentalist family by Cruft, but he would not concur with this classification (Cruft et al., 2015, p. 13). Despite broadening the array of goods protected by human rights to an indeterminate and open-ended list of things beneficial to humans (with the sole constraint that establishing a human right to X does not impose excessive burdens on corresponding duty-bearers), Tasioulas explicitly rejects the notion that his pluralist approach "dispenses with any reliance on the controversial notion of human dignity" (Tasioulas, 2015, p. 53). The 'ontological basis' of human dignity is that we are characterized by 'capacities for language-use, for registering a diverse range of normative considerations (including evaluative considerations, prudential, moral, aesthetic, and others besides), and for aligning one's judgments, emotions and actions with those considerations' (Tasioulas, 2015, p. 54).

[416] I do not believe that political approaches to human rights should be framed as alternatives to either instrumental or non-instrumental theories. In practice, depending on the author, so called "political accounts" have drawn on elements of both. They are also framed in opposition to moral (or so-called "orthodox") however, as already said, that this dichotomy is poorly formulated and ultimately unhelpful. Many scholars who identify with the political approach—such as Cohen (2006), Beitz (2009), and Ferrara (2014)—follow Rawls's methodology in Political Liberalism and appeal to public reason, especially in its global form, as the framework through which human rights can and should be identified and justified. Yet public reason, properly understood, is not merely a political tool—it is also, and primarily, a moral concept. Viewed in this light, these authors can be seen as pursuing a third path, one that parallels the constructivist approach I am highlighting—moving beyond the simplistic opposition between instrumental and non-instrumental justifications by integrating elements of both within a discursive and deliberative framework.

identified three main approaches: one emphasizes the notion of agency, another centers on the notion of the good life, and the third appeals to the notion of basic needs (Cruft et al., 2015, p. 11). All these different foundations share a common instrumental mode of justification, proceeding in two steps: identifying a valuable end and ascribing human rights to individuals because possessing these rights contributes to the realization of this end.

Agency is an obvious candidate for an account of human rights, since it is considered by many as the distinguishing feature of human beings. As we have seen, however, an account that appeals to agency, such that of Griffin, incurs in various problems.[417] One way to avoid some of this problem is to take a broader view according to which human rights are grounded in a plurality of goods that are required to have a good life, where agency is simply one of these goods. For example, John Finnis's list identifies life, knowledge, play, aesthetic experience, sociability (friendship), practical reasonableness, religion as the fundamental human good; on his account, human rights are justified to the extent in which they further these goods (Finnis, 2011).[418] Similarly, James Nickel adopts a pluralist perspective and argues that human rights are vindicated

[417] The primary problems lie in the exclusion of some characteristics that seems more relevant than agency in some examples such as torture and slavery. While the undermining of agency, specifically the capacity to decide and adhere to decisions, is a crucial aspect justifying a human right against torture, the violation of human rights, in these cases, is not solely or predominantly related to the deprivation of agency. Rather, we recognize that part of the wrongdoing involves the pain suffered by the victim, and thus absence of pain seems to be also a relevant feature. Furthermore, slaves possess a degree of agency, shaped by education and skills (essential for their value to a master); nonetheless, their human rights are still violated, indicating that agency alone cannot be the foundation of human rights. If Griffin's concept of agency encompasses more than the minimal education and resources provided to slaves, determining a non-arbitrary threshold for the goods sufficient to fulfill the requirements of agency becomes challenging. Intuitively, human rights are perceived not as rights to a good life but to a decent one. However, defining what decency requires for an entity endowed with agency poses a subjective challenge (Caranti, 2019a, p. 44).

[418] By asserting that there are certain human ends that are objectively good, Finnis means that their value doesn't derive from individuals merely desiring them but rather from their inherent status as fundamental components of human well-being. He refers to these as the "basic forms of human good" and constantly reminds that while discussing them he is not referring to any 'moral good' (Finnis, 2011, p. 86) but to what constitute the first principles of practical reasoning, which ground both the requirements of practical reasonableness and the foundations of justice. For Finnis, this framework enables a clearer articulation of legal obligation by distinguishing between intra-systemic legal obligation (formal validity within the legal system), and moral legal obligation (law's presumptive moral force). Intra-systemic legal obligations are often treated as binary and invariant—there being no degrees of legal obligation, just as there are no degrees of legal validity. This serves as a technical device for legal reasoning, particularly in adjudication and systemic problem-solving. It derives from authoritative stipulations under the system's rules, which function as *exclusionary reasons* for action (Finnis, 2011, pp.308-309, 354). Legal obligation in the moral sense, instead, refers to the moral obligation that is presumptively implied by intra-systemic legal obligation. This moral obligation arises for the conscientious citizen by integrating the legal stipulation into the unlimited flow of practical reasoning. It is based on the understanding that the legal system (or the specific law) provides directly applicable and authoritative guidance, eliminating the need for subjects to weigh all pros and cons, since the legislature (or competent authority) has already performed this evaluation for the "common good" (Finnis, 2011, p.351).

via ‘linkages arguments’, that is, showing that their exercise is necessary for, or at least significantly contributes to, meeting fundamental claims, which he sums to four: life, the steering of one’s life, avoidance of cruel and degrading treatment, avoidance of severely unfair treatment (Nickel, 2007, pp. 61–90)[419].

However, these perspectives appear to encounter the opposite challenge of the agency proposal, namely overinclusion. If any universal human interest that could be protected through the establishment of a corresponding human right (without imposing excessive burdens on duty bearers) is deemed to generate a corresponding human right, we may arrive at counter-intuitive results, given the widely shared intuition that human rights are meant to establish minimum standards rather than to promote the good life[420]. For instance, if individuals could avoid workplace boredom (perhaps through the rotation of monotonous tasks), and if a social system could be devised where this can be achieved without imposing excessively high costs on duty bearers, then there would seemingly be a human right not to experience boredom in the workplace (Caranti, 2019a, p. 46).

To avoid the underinclusion resulting from focusing solely on normative agency and the overinclusion problem of goods, some philosophers have moved to a middle-grounded idea: that we can ground human rights by appealing to a particular class of human needs. Although typically the needs we have depend on the adoption of specific goals (“I need to study if I want to become a professor”) one can argue that there is a class of needs that all human beings have simply qua human beings, that do not seem to depend on the adoption of any specific goal or plan of life. Henry Shue for example posits that there exists a set of basic needs inherent to all humans, independent of their conception of the good life; these needs encompass essentials like food, water, and air, as well as a minimal level of social interaction and recognition (Shue, 2020).

While this approach may avoid the underinclusion associated with the agency-based approach and the overinclusion of the second approach, it encounters at least two challenges. Firstly, determining the 'golden mean' appears somewhat arbitrary: who decides how basic a need

---

[419] Nickel exemplifies the limitations of these subdivisions, since he is speaking of claims and thus his perspective aligns with both the good life and the basic need approaches potentially.

[420] As Shue remarks, human rights seems to be more related with ‘the lower limits on tolerable human conduct’, rather than ‘great aspirations and exalted ideals’ (Shue, 2020).

must be to warrant inclusion among those generating human rights? Secondly, quintessential human rights such as first-generation civil and political rights can be traced back to basic human needs only by some stretching. Consider the right to equality before the law or the right to fair trial. Which basic needs do these rights aim to protect exactly? Defenders of the basic needs approach argue that these rights can be vindicated via 'linkages arguments', namely by showing that their exercise is necessary for, or at least significantly contributes to, meeting fundamental needs such as bodily security or subsistence, but opponents of this view worry that this link might be too thin and indirect to provide a suitable justification for civil and political rights.

A response to this criticism may be offered from the capability approach developed by Amartya Sen and Martha Nussbaum. This framework, arguably the most refined conception within the needs subgroup, establishes a direct link between human rights and capabilities—defined as the actual opportunities individuals have to realize their life plans (Nussbaum, 1997, 2000; Sen, 1984, 2004). Nussbaum explicitly lists ten basic human capabilities (life, bodily health and integrity, senses, imagination and thought, emotions, practical reason, affiliation, relations to other species, play and control over one's environment) and thinks that all human beings are entitled to these things as a matter of (human) rights.

Despite its ambitious goals, the Capabilities Approach has inherited some of the challenges faced by others instrumental approach, like under-inclusion; similarly to Griffin's agency-based approach, the Capabilities Approach struggles to account for certain children's rights. Many of these rights—such as the right to health care, education, a name, and nationality—are centered on children's current functionings rather than their capabilities. Nussbaum attempts to address this by arguing that these functionings are crucial for developing adult capabilities. However, this response overlooks the problem of potentiality, [421] and the reality that some children, such as those with terminal illnesses, may not reach adulthood. Additionally, the Capabilities Approach seems to face an over-inclusion problem, with some scholar arguing that Nussbaum's list of essential capabilities may include too much (Freeman, 2006). For instance, while being loved is undoubtedly one of life's most important experiences, does it make sense to frame it as a right?

[421] This is exemplified in the discussion regarding fetal rights. As John Harris has pointed out, the potentiality argument proves too much: if the embryo is potentially a person, then it is also potentially dead, like any of us. The possibility of an organism acquiring certain characteristics in the future that it does not possess in its present state does not justify us treating it as if it has already acquired them (Harris, 2001).

As Joseph Raz points out, the significance of something to human well-being does not alone suffice to establish it as a right (Raz, 2010a).

This critique highlights a general issue with instrumental justifications: while they underscore the importance of certain aspects of human life, they do not, by themselves, justify rights claims. Instead, they reinforce the value of particular goods or capabilities without establishing an entitlement to them (Caranti, 2019a, p. 47).[422] In order to avoid both being under-inclusive or being over-inclusive, each of the positions mentioned so far has to draw on some distinction between fundamental and non-fundamental aspects of agency, of the good life or of needs: not all needs generate rights, nor all aspects of the good life or of agency. Drawing this fundamental/non-fundamental distinction in a substantive, non-question-begging way is a challenge for each account.

3.3.2 The Orthodox View: Human Worth and Status

According to the Orthodox (or non-instrumentalist) view, human rights are universal moral rights possessed by all human beings simply by virtue of their human status.[423] They don't offer instrumental justifications for human rights but argue that we hold human rights as a matter of our basic moral status, and that our holding these rights is at least partially independent of whether and how they promote or protect further human values such as agency, needs, freedoms, or interests. While instrumentalism starts with the identification of some more or less objective human goods and then moves to their transformation into things we have rights to, non-instrumentalism relies on the worth human beings supposedly have. In this sense, the point of human rights is not to protect the interests of humans (the parts of their being that they value) but to protect a status humans possess (Caranti, 2019a, p. 48).

[422] "To put it crudely, instrumental justifications do not justify any right: they simply reinforce the importance for human lives of certain things, with variations as to what those things are. They hardly provide an argument to move from the importance of the goods to the claim that we are entitled to them." (Caranti, 2019a, p. 47).

[423] Tasioulas takes the term from Beitz, that contrast the 'orthodox' view with a 'practical' conceptions of human rights (Beitz, 2004, p. 196).

Allen Buchanan (2010) in arguing why human rights cannot be grounded solely by appealing to the goods they secure and protect, provides a very illuminating example: a female chief executive, well paid for the top position she occupies, can still be wronged if she is paid slightly less than a male top executive merely because she is a woman. Even if none of the goods the woman needs for a decent or even good life are put at risk by the lower compensation, it still seems that we have here a clear violation of human rights. This, according to Buchanan, indicates how human rights have to do with the protection of a status, of an equal status to be precise, quite independently of whether and how they promote human well-being. In this sense, a deontological element seems to be irreducible in any convincing account of rights in general and of human rights in particular. Any proposal that fails to adequately address this aspect is likely to be unconvincing or incomplete, since: "The fact that strong rights against discrimination are prominent in a convention devoted to the special problems of discrimination faced by women, along with the fact that there is growing attention to issues of gender discrimination on the part of various nongovernmental international human rights, is evidence that these rights are a feature of IHR that a philosophical reconstruction must acknowledge" (Buchanan, 2010, p. 687).[424]

Frances Myrna Kamm, in a broader argument discussing the moral importance of some non-consequentialist distinctions, endorses this deontological assumption in the human rights debate by presenting some notions of moral status that are based on the intrinsic worth of the person: "fundamental human rights […] are not concerned with protecting a person's interests, but

[424] The gender wage gap, while a clear example of injustice, is a complex and contentious issue regarding its origins and causes. Some scholars attempt to justify the gap based on biological differences, for example that women are less likely than men to negotiate for salary increases, potentially influenced by hormonal factors like testosterone. Nobel Prize winner Claudia Goldin's extensive research, spanning over two centuries of labor market data, has tried to uncover the deep historical roots of gender wage disparities. Her work emphasizes how societal changes—such as the advent of the contraceptive pill, which empowered women with greater control over family planning, and the growing acceptance of remote work—have gradually improved women's earning potential. These trends indicate that further progress, including flexible work arrangements during maternity leave, could help us achieve greater wage parity (Goldin, 1990). Nevertheless, this debate does not undermine Buchanan's argument that human rights must ensure egalitarian status for all individuals rather than focus solely on specific characteristics. His argument specifically addresses cases where women are denied equal pay for equal work, even when they would want to negotiate for it. Such situations represent clear injustices and diminish their status, regardless of any claims related to maternity leave or similar factors. This principle is widely recognized, as pay discrimination based on gender is illegal in the majority of countries worldwide. However, the need for a comprehensive framework for human rights is particularly urgent in countries where such protections are not fully legalized or enforced, such as Saudi Arabia and Afghanistan.

with expressing his nature as a being of a certain sort, one whose interests are worth protecting. They express the worth of the person rather than the worth of what is in the interests of that person (Kamm, 2006, p. 271). “On this account, it is impermissible for me to harm the person in order to save many in the accident, because doing so is inconsistent with his having this status”(Kamm, 2006, p. 254).

Thomas Nagel (1995) also sees human worth as the fundamental basis of human rights, and, like Kamm, he does not find it necessary to justify that status. Nagel halts any further inquiry by arguing that “any attempt to render more intelligible a fundamental moral idea will inevitably consist in looking at the same thing in a different way, rather than in deriving it from another idea which seems at the outset completely independent” (Nagel, 1995, p. 92). The view defended by Kamm and Nagel is non-instrumental in that human rights are not justified by appealing to their capacity to promote, protect or further certain valuable ends, but rather by appealing to their role in expressing or reflecting the moral worth of their holders.

This, however, is not the only way to justify human rights in a non-instrumental fashion. According to a different approach, which might be termed ‘transcendental’, human rights can be justified by appealing to the idea that to be an agent at all, one must conceive oneself and others as bearing human rights. Alan Gewirth 'transcendental' argument starts precisely from the idea that we are all purposive agents (we all act for some good that we consider worth pursuing) and since freedom and well-being are necessary conditions for the pursuit of any end, it follows that we are constrained to want freedom and well-being: this is equivalent to saying that we want the freedom and well-being of all to be protected through human rights (Gewirth, 1984).

The primary criticism of non-instrumental approaches is that, without a clear understanding of what truly matters to humans, and relying solely on the bare notion of possessing a status valuable per se, orthodox thinkers may fail to adequately address the diverse areas of human interest that human rights encompass, which range from civil and political liberties, to more fundamental prerequisites such access to food, shelter, and bodily integrity (Nickel, 2007; Tasioulas, 2015). However, adopting a non-instrumental perspective does not necessarily entail rejecting the idea that human rights are grounded in the plurality of goods they aim to protect: human rights can be founded both on the inherent status of individuals and on their capacity to promote, further, or safeguard specific objective human goods. (Cruft et al., 2015, pp. 17–18). Still

some scholars, such as Caranti, argue against this syncretic approach that attempts to combine instrumentalism and deontology, as it risks conflating two distinct levels of foundation that should be clearly differentiated (Caranti, 2019a, p. 49).[425]

3.3.3 A Relational Foundation Beyond Instrumentalism and Non-Instrumentalism

Thus, if non-instrumentalist is the framework according to which human rights are not justified by appealing to their capacity to promote, protect, or further certain valuable ends (bottom-up), but by appealing to their role in expressing or reflecting the moral worth of their holders (top-down), it seems that all scholars that want to maintain a conception of dignity, even if as a status-concept and not a value-concept, are deemed to fall inside the non-instrumentalist orthodox family with all the problem attached to them, such as issues of applicability and feasibility. This can be avoided by either refuting or appropriately specifying in which sense dignity is a foundation for the concept of human rights.

In a specific paper, called precisely "Is Dignity the Foundation of Human Rights?", Waldron directly address in which sense he takes is concept of status-dignity as a "foundation": he is referring neither to an ontological basis nor to a logical derivation of those rights, but to a concept that allows us to better grasp the significance of human rights (Waldron 2015, p.11). He explains that there are four ways in which one concept (such as human dignity) might be the foundation of another (such as human rights): the foundational relationship can be understood in terms of (i) one concept arising from another as a matter of "history and geneology", (ii) one serving as the source of the validity or legitimacy of the other, (iii) one being logically derivable from the other (perhaps with the addition of certain empirical premises), or (iv) one concept throwing "some indispensable light" on the other (Waldron, 2015, p. 12). The third and fourth notions of a foundation are the more significant in the contemporary debate, with the third representing what philosophers and legal or political theorists no doubt chiefly have in mind when

[425] "On one level, there is the question of what provides normative force to the idea that humans have a status that entitles them to certain basic rights. Let us call this the foundation stricto sensu. On a less profound level, there is the question of what goods need to be secured for having a life compatible with that status. The second is mainly a technical question that depends on the level of our social and psychological knowledge, on historical circumstances, on the average welfare of a given society and on many other variables that should not be made part of a normative justification. In short, the second question, despite appearances, is not part of a foundation at all" (Caranti, 2019a, p. 49).

they speak of foundations,[426] even if the fourth, in Waldron's view, more accurately captures the relationship between human dignity and human rights.

Constructivists too consider dignity as a constructed concept, not a foundational one in the other third sense. Rather than conceptualizing dignity as the 'foundation' of human rights, constructivists insist on the 'dynamic' nature of dignity as a relational status requiring respect. The difference with Waldron is that for a constructivist this status is not socially constituted, but rationally constructed, and as such it maintains a normative orientating force that is less stable in socially contingent approaches like Waldron's. The normativity of the status is granted not through a foundational (logical or ontological) grounding: it arises from its constitutive authority in relation to my own first-person perspective and the public standpoint of second-person relationships.[427]

Constructivists, who reject the conventional strict opposition between utilitarianism and deontology (Bagnoli, 2016), will also contend that the distinctions between instrumental and non-instrumental justifications of human rights become less significant. Human rights are neither merely instrumental (as tools) nor solely non-instrumental (as intrinsically worth); instead, they are constructed standards that articulate the appropriate way to regard and treat other persons within a moral community. This constructed nature of rights allows for a nuanced, pluralistic position: human rights can be understood as having both a non-instrumental value—rooted in the relational status of dignity—and an instrumental role in advancing other important values, such as personal agency, social participation, or the ability to live a minimally decent life. This pluralist constructivist position allows for human rights to be justified not only on the basis of moral worth but also in terms of their practical role to act appropriately in the contemporary world.

For Forst, for example, human rights are not justified instrumentally, but in the rational acknowledgment of ourselves and others as beings capable of justification. It is through this mutual recognition that we acknowledge one another as authorities within the "space of reasons"—individuals who possess the ability to engage in and demand justifications. Only after this

[426] Waldron takes James Griffin's theory of human rights to be an example of this third approach to showing that human dignity is the foundation for human rights, with Griffin's ideas about the importance of "normative agency" (and about what normative agency involves) constituting the relevant analytical unpacking of the concept of human dignity. Waldron have no particular interest to reject foundational derivations of his third sort, Griffin's or any other, but he is still convinced that understanding the point of (human) rights in such "a rigidly teleological sense" is less fruitful than the fourth approach (Waldron, 2015, p. 19).

[427] We will return to this point in the next chapter, specifically paragraph 4.2.1.

recognition, the conception of dignity is introduced as a constructed foundation of a "community of justification" where each person is respected as an equal participant who owe each other good reasons for how they treat one another and for the normative order to which they are subject (Forst, 2012, p. 14).[428] In this sense, dignity is not simply a static attribute, but a dynamic status that emerges from the ongoing process of reason-giving and justification. As he puts it: [T]he notion of "dignity" that lies at the heart of such a conception of human rights is not a metaphysical or ethical one, combined with a doctrine about the good life. Rather, dignity means that a person is to be respected as someone who is worthy of being given adequate reasons for actions or norms that affect him or her in a relevant way. And this kind of respect requires us to regard others as autonomous sources of normative claims within a justificatory practice. Each person is an "authority" in the space of reasons, so to speak. Dignity is thus a relational term; its concrete implications can be ascertained only by way of discursive justification" (Forst, 2010, p. 734).

Dignity, therefore, serves as the overarching term that encapsulates the status attributed to every individual as a rational and autonomous being we recognize as entitled to respect. This notions of rationality, autonomy and respect are also strictly connected with accountability and responsibility: "a person acts autonomously-that is, as a self-determining being when she acts intentionally and on the basis of reasons. She is aware of the reasons for her action, can "respond" when asked for her reasons, and is thus "responsible" for herself. Autonomous persons in this sense are accountable agents, accountable for themselves to both themselves and others; they can reasonably explain and justify their actions." (Forst, 2012, p. 129). Contrary to some widespread views, autonomy is not a rejection of our relational and dependent nature but a practical stance we adopt to become authors of our own actions that implies relations with others (Bagnoli, 2007, p. 125).[429] The notion of dignity, when understood as a status in this qualified sense, captures all the

---

[428] "The ground created by reasons must therefore be a shared, common basis for justified, well-founded thought and action. 'Standing their ground' means that the things being justified can withstand challenges and relevant criticism, and that the respective reasons thereby become common property: reasons may refer to very specific beliefs and actions, but as reasons they are in principle publicly accessible. They can be 'provided,' 'accepted,' and 'demanded,' and they are not private property. They can be generally assessed according to criteria of reason and are part of the common 'game of giving and asking for reasons'." (Forst, 2012, p. 14).

[429] "The fortune of this objection about the limited competence of respect is partly due to a misunderstanding of the conception of autonomy. Contrary to what the objection implies, autonomy is not a denial of our dependent and relational nature. More generally, autonomy is not a hypothesis we concoct about our ontological status, but a practical stance we undertake toward ourselves in order to acquire authorship on action. It is a practical response to the many ways we are dependent, rooted, and situated. To represent ourselves as autonomous means to represent ourselves as

ideas of autonomous respect and responsible mutual accountability, without reverting to a strictly non-instrumental conception, as it functions not as an intrinsic foundation for human rights, but as a relational concept that conveys all these normative ideas within a constructivist (non-realist) framework.

3.4 FROM HUMAN RIGHTS TO PERSON RIGHTS?

This constructivist reconceptualization of dignity as a status attributed to beings deserving of respect raises one final, pivotal problem—the very question with which we began: who qualifies as the beneficiaries of this status? If dignity is tied to the capacity for reasoning and the ability to participate in justificatory practices, are we then excluding human beings who have lost, or never possessed, the capacity for rational thought and moral reasoning? This would include individuals with severe cognitive disabilities, advanced dementia, and even infants, who cannot actively engage in demanding justifications or reciprocate in the moral community in the same way. Do these individuals still possess human rights if they lack the status to demand them?

Scholars who answer negatively to this final question are often labeled as "rationalist elitists" (Nino, 1991, p. 36). Rather than justifying "the existence of rights that all human beings possess qua human being, they justify something different: rights that all persons possess qua persons" (Cruft et al., 2015, p. 12). James Griffin for example notes that human rights are protections of "our normative agency" or of an "essential feature of personhood" (Griffin, 2010, p. 345f) grounded in the two basic human interests of autonomy and liberty (Griffin, 2008, p. 32). Similarly, Carl Wellman submits that "one can and should accept the conclusion that only human beings defined not as members of a biological species but in the morally relevant sense as persons with the normal human capacities for moral action could possess any moral human right."

capable of deliberating about issues that arise exactly because of our relational and dependent nature." (Bagnoli, 2007, p. 125).

(Wellman, 2011b, p. 22).[430] Forst's constructivist definition, which centers on justifying beings, also falls within this category (Forst, 1999, p. 43).

This focus on the concept of person over that of human beings, have brought some scholars to discuss of person's rights (Buchanan, 2011, p. 213 fn).[431] The term "person" in this context is not simply synonymous with "human being"; rather, it is specifically applied to subjects who possess certain defining characteristics or capacities, such as rationality, autonomy, moral agency, or a sense of self.[432] These accounts are often regarded as "morally reprehensible" in their restrictive stances (G. Quinn & Arstein-Kerslake, 2012); since such perspectives undermine the protective scope of human rights (excluding some of society's most vulnerable and non-paradigmatic members) and conflict with our common and legal understanding of universal human rights (rights that every human being holds, regardless of their individual characteristics, including physical or mental abilities).

Constructivists typically seek to distance themselves from the more extreme forms of rationalist elitism. O'Neill, for example, argues that we should respect rational nature in the abstract, not only as it is realized in fully rational persons. This entails respecting the fragments or necessary conditions of rational nature embodied by individuals—such as small children or those with severe mental impairments—even when full rational capacity is absent (Wood & O'Neill,

[430] "It may seem monstrous to deny that infants can have any moral human rights. But no moral philosophers, and certainly no parent, imagines that a suckling babe has a moral duty not to inflict pain upon its mother by biting her breast or that a very young child has a moral obligation to refrain from throwing disliked food onto the floor. We all realize that children acquire moral obligations gradually, as they grow to become responsible moral agent. Why not recognize that similarly infants acquire their moral rights only as they develop their capacities for moral agency?" (Wellman, 2011b, p. 22).

[431] " If "humanity" refers to personhood [...] then we might decide that what we have called human rights would be more accurately called persons' rights.' (Buchanan, 2011, p. 213 fn).

[432] Since at least John Locke, the concept of personhood has been linked to the capacity for "second-order" attitudes—such as the ability to approve or disapprove of one's own desires—which render them subject to normative evaluation: while I may desire to act in a certain way, I might simultaneously disapprove of that desire and choose to resist its influence, with my disapproval itself motivating me to refrain from acting on it. A more recent formulation of this idea comes from Harry Frankfurt (1988), who argues that second-order attitudes are second-order desires—desires about which of our first-order desires we should endorse and act upon, and which we should reject. According to Frankfurt (1988, pp. 11–25) personhood is defined by the capacity to form and be guided by such second-order desires, distinguishing persons from beings who merely act on their immediate impulses.

1998, p. 198).[433] She emphasizes that, although these individuals may not qualify as "persons" in the strict technical sense, they are nonetheless owed moral respect.

However, there are some scholars who conceive the concept of person as wide enough to still include all human beings. Gewirth's approach to human rights for example, while still based on the idea that rights are tied to agency, interprets agency broadly enough to include all human beings (Gewirth, 1980, 1984). Even those who may not currently exercise full rational agency like infants or individuals with disabilities, are considered "prospective purposive agent" (Gewirth, 1984, p. 10) and thus are entitled to rights. Similarly, Nussbaum's capabilities approach provides a much more inclusive framework: rather than limiting personhood to rationality or autonomy, her approach emphasizes the full range of capabilities necessary for living a flourished life (Nussbaum, 1997, p. 285). One of the key strengths of this proposal is its inclusivity, as it extends to all human beings, including those with profound disabilities who may lack certain specific abilities: the focus is on what individuals are able to do and be.

Nussbaum framework, however, extends even further. In her book *Frontiers of Justice* (2007), where she criticized contractarian theories of justice for their inherent exclusion of beings who cannot participate in reciprocal agreements, she further argued that this exclusion not only affects individuals with severe disabilities and children but also extends to non-human animals. Nussbaum argues that animals, like humans, should be entitled to develop and exercise their natural capabilities: this entails providing them with the necessary conditions to thrive in ways that align with their species-specific flourishing. Animals should not be treated merely as resources for human use but as beings with their own moral entitlements. By integrating animals

---

[433] "Of course we should respect rational nature in persons, and this means respecting the persons themselves. But my main argument here depends on saying that we should also respect rational nature in the abstract, which entails respecting fragments of it or necessary conditions of it, even where these are not found in fully rational beings or persons. The point I am making is easiest to see, and hardest to deny, in the case of many human beings (in the nontechnical sense) who lack 'humanity' (in the technical sense), and therefore must fail (technically) to be persons at all. They include small children and people who have severe mental impairments or diseases which deprive them, either temporarily or permanently, of the capacity to set ends according to reason. Clearly Kant would not want to say that such human beings are mere things, which are to be treated only as means. The important thing, though, is not what Kant would want to say, but rather what is required by a reasonable interpretation of his basic principle that rational nature should be respected as an end in itself. The point is that it would show contempt for rational nature to be indifferent to its potentiality in children, and to treat children as mere things or as mere means to the ends of those beings in whom rational nature is presently actual. Owing to the fragility and vulnerability of the potentiality for rational nature in children, Kant's principle might even dictate giving priority to its development in children over promoting some of the ends of actual rational beings." (Wood & O'Neill, 1998, p. 198).

into her framework, Nussbaum not only broadens the scope of moral consideration but also underscores the interconnectedness of human and non-human flourishing, advocating for a more compassionate and inclusive vision of justice (Nussbaum, 2007, pp. 399–401).

The question of animal rights and its relationship to human rights has become a crucial and contemporary debate. A key challenge against rationalist accounts of rights is their inconsistency when addressing marginal cases or species overlap. If some humans and animals possess similar cognitive abilities, the principle of equal treatment forces a dilemma: either exclude humans who lack rationality alongside animals (Wellman, 2011b, p. 22),[434] or extend rights to animals with comparable cognitive capacities (Kagan, 2016, p. 6). Several constructivist scholars have directly engaged with these questions. For example, Korsgaard addresses the issue in her influential paper "Fellow Creatures: Kantian Ethics and Our Duties to Animals" (2004), which later evolved into the full book *Fellow Creatures: Our Obligations to the Others Animals* (2018).

In this work, Korsgaard, drawing on Kant, defines persons as rational beings who can self-legislate and participate in the moral community through the mutual imposition of obligations and she believes that this definition does not leave out any human beings; animals, however, lacking this capacity for self-legislation, do not fall into this category of persons (Korsgaard, 2018, pp. 97–98).[435] Nonetheless, Korsgaard reject Kant's view that our duties to animals are "indirect," that is, owed to ourselves rather than to the animals, since they can still be considered ends in themselves,

[434] "Therefore, a necessary condition for the possession of any moral human right is the capacity for moral action of any normal adult human being. But some biological human beings, especially neonates and the irreversibly comatose, are incapable of acting in any morally relevant sense. Therefore, it is only beings that possess the normal human capacities for acting who could possibly possess any moral human right. It may seem monstrous to deny that infants can have any moral human right. But no moral philosopher, and certainly no parent, imagines that a sucking babe has a moral duty not to inflict pain upon its mother by biting her breast or that a very young child has a moral obligation to refrain from throwing disliked food onto the floor. We all realize that children acquire moral obligations gradually as they grow to become responsible moral agents. Why not recognize that similarly infants acquire their moral rights only as they develop their capacities for moral agency?" (Wellman, 2011b, p. 22).

[435] "I argued that when the ideas of rationality and moral standing are properly understood, Kant's account does not leave out any human beings. All human beings are constructed to function rationally, although some may have defects that make that difficult or impossible for them. There is a difference between being a defectively rational creature, and being a different kind of creature altogether. But although we should keep in mind that it is an empirical question, it seems likely that only human beings are rational. So that part of the question still stands: If we human beings are ends in ourselves because we are rational, then do the other animals lack moral standing? Kant himself thought so […] Animals are among the beings whose existences "rest on nature," and so here are declared to be "things" which may be used as mere means." (Korsgaard, 2018, pp. 97–98). To further explore the relationship between Kant and animals, refer to the book edited by Callanan and Allais (2020).

in a passive sense: while they are not capable of self-legislation like rational beings, animals have something that is good for them (Korsgaard, 2018, p. 141).[436] This "final good" (Korsgaard, 2018, pp. 156–165) is based on their ability to be sentient, namely their capacity to experience pleasure and pain and their evaluative responses to what happens (Korsgaard, 2018, p. 20).[437] It is on this good that animals have that we should values them, not merely as means to our ends, and it is this good of animals that must be respected.

Korsgaard, therefore, acknowledges that while the respect we owe to animals differs from that which we owe to persons, it is undoubtedly morally significant. While obligations toward persons stem from our capacity to be members of a moral community that mutually self-legislates, obligations toward animals, on the other hand, arise from the recognition that they too have a final good valuable for them derived from their nature as sentient beings. "[W]e obligate each other by making claims on each other that are reasonable because in making them we allow that others may reasonably make similar claims on us. The other animals obligate us by reminding us of what we as individuals have in common with them—that we are creatures for whom things can be good or

[436] "In fact, there are two slightly different senses of "end in itself" at work in Kant's argument, which we might think of as an active and a passive sense. I must regard you as an end in itself in the active sense if I regard you as capable of legislating for me, and so as capable of placing me under an obligation both to respect your choices, and to limit my own choices to things compatible with your value as an end in itself. That is what it means to see us as bound together in a system of reciprocal legislation. I must regard you as an end in itself in the passive sense if I am obligated to treat your ends, or at least the things that are good-for you, as good absolutely. Kant evidently thought that these two senses come to the same thing: that everyone who is an end in itself in the passive sense must be an end in itself in the active sense too. When you are just thinking about people, the two ideas do seem very close, for to demand that your end should be treated as absolutely good is just to demand that everyone should treat it as giving them reasons, and to do that is just to claim that you have the standing to make a law for others, to participate in reciprocal legislation. But when we think about animals, it becomes obvious that the two ideas are not the same. It certainly could be true that animals are ends in themselves in the sense that what is good for them is good absolutely, even if they are not capable of joining with us in reciprocal legislation." (Korsgaard, 2018, p. 141). This emphasis on the passive stance is part of a broader discussion—one that Korsgaard (2018, p. 189) briefly touches on—regarding the distinction between moral agents and moral patients. This distinction is also highlighted by Richardson in his definition of the members of the moral community. We will revisit this point at the end of this chapter.

[437] This is the general definition of non-rational animal that Korsgaard gives: an organism that interacts with its environment by representing it through sensory perception and acting based on those representations. Furthermore, to function effectively, an animal must distinguish between what is good or bad for her in a functional sense. This requires a "valenced" representational system, where environmental stimuli appear as either attractive or aversive, pleasant or painful. As Korsgaard puts it: "In order for an animal's representational system to do its work in this way, however, it has to have what I will call a "valenced" character. That means that the things she encounters in her environment have to strike her as attractive or aversive, welcome or unwelcome, pleasant or painful, in particular ways, depending on whether and how they are good- or bad-for her. She has to be drawn by the way things appear to her to seek out the things that are good-for her and to avoid the things that are bad. So she has to perceive the world evaluatively, as a place full of things which present themselves as attractive and to-be-sought and things which present themselves as aversive and to-be-avoided." (Korsgaard, 2018, p. 20).

bad, and that like them, although in our own special way, we each take our own good to be good absolutely when we engage in practical activity." (Korsgaard, 2018, p. 147). By recognizing this distinct yet equally important mutual relationship, we understand that our obligations toward animals are not merely duties we owe to ourselves, but rather reflect a correlative relationship between their right to be free from suffering and our duty to alleviate their pain and refrain from causing it unnecessarily.

Korsgaard's solution has been criticized, especially for trying to remain inside a Kantian framework that, has admitted also by herself, does not view animals as end in themselves. The leap from Kant's "Formula of Humanity"—which grounds moral standing in rational self-legislation—to what has been described as a "Formula of Animality" (Darwall, 2022, pp. 213–214), has struck many Kantian scholars as theoretically strained and insufficiently motivated. Darwall (2022a) analyses in detail how Korsgaard's argument, while effectively elevating the moral status of animals, ultimately fails to establish them as bearers of rights in the Kantian sense, which is fundamentally tied to agency and mutual accountability.[438] However, even Darwall suggests that an alternative, accountability-centered approach—one that reframes obligations to animals in terms of our responsibilities rather than their capacity for a good—could offer a stronger foundation for extending moral consideration to animals within a Kantian framework (Darwall, 2022a, pp. 235–236). If we accept that rational agents have a duty to recognize and protect the vulnerable, including those who lack the capacity for moral accountability, we may be able to justify moral obligations toward animals not merely on the basis of their suffering but as part of our broader moral responsibility toward them. This would require expanding the concept of moral standing to include beings who cannot participate in moral accountability but who nevertheless deserve protection due to their inherent vulnerability and dependence.

Still, some scholars argue that even such "moderate rationalist" positions—by drawing a clearer distinction between persons and animals—miss the broader point. For instance, Saskia Stucki (2023) contends that the notion of personhood risks systematically excluding animals from

[438] The demand theory of rights has produced several important works related to the topic of animal rights. Feinberg, particularly in his essay "The Rights of Animals and Unborn Generations," explicitly defends the idea that beings who are incapable of claiming their rights (e.g., infants or those with severe cognitive impairments) can still possess rights and argues for a broader moral community that does not exclude humans based on a lack of rationality or autonomy (Feinberg, 1974).

the full sphere of justice, thereby reinforcing an anthropocentric framework that fails to account for their moral and legal status, creating an artificial boundary that overlooks the shared experiences of suffering, vulnerability, and moral consideration between humans and non-human animals. In response, Stucki advocates for a "One Rights" theory, in which human and animal rights are not seen as two separate categories but as part of a unified system of rights based on shared interests, vulnerabilities, and experiences of injustice.[439] Within this framework, fundamental rights such as the right to life, liberty, and physical integrity could be recognized not only for human beings but also for animals, thereby challenging the traditional boundaries of moral and legal consideration (Stucki, 2023, pp. 95–96).[440]

While, therefore, Saskia Stucki argues that the notion of personhood is just as misleading and exclusionary as the concept of "human" when it comes to rights (Stucki, 2023, p. 22),[441] some scholars believe that personhood can be retained without excluding animal rights (Francione, 2006).[442] One could say that, following the two shifts Pogge identifies from natural law to natural

---

[439] It is interesting to note how Stucki use a three stage process similar to the Constructive Ethical Pragmatism of Richardson, although she take it from Hurst Hannum (2016). "At present, animal rights are new human rights claims and as such 'merely candidates for legal recognition'. Broadly speaking, we can distinguish three stages in the 'birth process' of a new human right: from its intellectual inception (the idea phase), to its gradual reception and consolidation in legal and political arenas (the emergence phase), to its eventual legal recognition and codification (the recognition phase). Animal rights are currently located in between the first and second stage of this 'lengthy period of gestation'." (Stucki, 2023, p. 4).

[440] "With regard to the practical dimension of One Rights, the review of political justifications has highlighted the practical interdependence of human and animal rights. Humans and animals—being part of the same planetary community, sharing many of the same environments, and often living in the same societies and political communities—are naturally and socially interdependent. It seems reasonable to think of human and animal rights, as normative protections of the natural and social preconditions for the enjoyment of rights, as equally interdependent. […] One Rights marks a clear departure from the terminologically reinforced truism that human rights are (exclusively) human, and submits that some human rights are animal rights. That is, some of the fundamental rights that humans have are the same kind of fundamental rights that animals (ought to) have, in virtue of their shared rights-generative properties. In terms of the specification of animal rights, there is a range of existing (civil, political, and social) human rights that may be extended and adapted to animals, to the extent that animals display the prerequisite rights-generative interests, needs, vulnerabilities, or experiences of injustice." (Stucki, 2023, pp. 95–96).

[441] "Grounding human rights in personhood may thus further complicate, rather than resolve, the matter of finding a coherent justificatory basis for human rights." (Stucki, 2023, p. 22).

[442] "The exploitation of the nonhuman great apes is immoral for the same reason that is immoral to exploit the hundreds of millions of mice and rats who are routinely exploited in laboratories or the billions of nonhumans who we kill and eat: the nonhuman great apes and all of these other nonhumans are, like us, *sentient*. They are conscious; they are subjectively aware; they have interests; they can suffer. No characteristic other than sentience is required for personhood." (Francione, 2006).

rights and from natural rights to human rights (Pogge, 2008, pp. 60–63)[443], it is now time to consider a third shift: from human rights to persons' rights, in the broader conception of person that is able to include also non-human sentient animals.[444] Under this perspective, a general right to be free from unnecessary suffering would apply universally to all persons—whether they are adults, children, or non-human animals—based on their shared capacity to feel pain. This move does not seek to reduce human rights to persons' rights, but rather to acknowledge that some rights are more specific to human beings, while others apply more broadly to all sentient persons. For example, the right to religious freedom remains exclusively a human right, as non-human animals lack the cognitive capacities necessary for religious belief or practice. We should, therefore, conceptualize these rights as forming nested subsets, much like an Eulerian diagram. Certain rights—such as the right to life or the right to health—belong to the broader category of persons rights, encompassing all sentient beings. Meanwhile, distinctly human rights, such as political

[443] Pogge highlights that the transition from natural law to natural rights (the first shift) focused on the rights-holder, emphasizing that violating a right primarily harms an individual, not merely a law. Natural law lacks this perspective—it does not necessarily impose constraints on a person's conduct toward others, and even if it does, violating those constraints does not necessarily constitute a wrong against individuals. Such violations might instead be considered offenses against God or disruptions of the harmonious cosmic order. This shift narrowed the scope of moral concerns that could be framed in terms of rights (Pogge, 2008, p. 61). For example, religious duties, duties toward oneself, or moral obligations toward animals are harder to articulate in the language of rights (the potential "third shift" would also possible involve reclaiming the animal concerns if the concept of persons is defined enough widely). The second shift, from natural rights to human rights, introduced four significant developments according to Pogge: 1) Secularization and focus on rights language. Human rights diverged from their medieval Christian roots, emphasizing the secularization process that began with the shift from the language of laws (commands and duties) to the language of rights. This evolution reinforced the connection between a special class of moral concerns and the status of rights-holders as subjects of moral value. 2) A political rather than metaphysical orientation. The term "human," unlike "natural," avoids implying an ontological status independent of human efforts or decisions. These framing steers clear of metaphysical and metaethical debates, making human rights accessible even to those who reject moral realism. 3) Equality of human beings: Human rights apply equally to all and only to human beings, who are regarded as equal in this sense. Every human being possesses the same rights, and their violation carries the same moral weight for all. 4) Protection from official threats: Human rights specifically safeguard individuals from violations by particular sources, especially "official" ones. Such violations are typically perpetrated by governments, government agencies, officials, or, to a lesser extent, other powerful entities (Pogge, 2008, pp. 62–63). These four points illustrate how the concept of human rights evolved from natural rights, introducing a more secular, politically oriented framework centered on institutional accountability. We will return on this evolution in the next chapter.

[444] Some argue for an even broader concept of personhood—one that extends not only to non-human animals but also to other natural entities. This idea is not merely an abstract philosophical debate: it has already found application in legal frameworks. In March 2017, the Whanganui River in Aotearoa New Zealand became the first river to be granted legal personhood. This status is rooted in an ontological understanding of the river as an indivisible, living whole and as the spiritual ancestor of the Whanganui Iwi, a Māori tribe (Kramm, 2020). Whether the inclusion of such entities should be considered part of the third shift in personhood, or perhaps a distinct fourth shift, remains a topic of debate. Here, however, I will focus primarily on the potential expansion of personhood to include non-human animals—an already highly contested move within Kantian circles.

participation or freedom of religion, exist within a more specific subset, applicable only to human persons.

While some constructivist proposals may not align well with this project (Forst, 1999, 2010), some, like those put forward by Bagnoli (2007, 2022) and Richardson (Richardson, 1994, 2018), offer a more flexible and broad conception of the moral community and the rights-holders within it, accommodate various agents without prematurely narrowing the scope of inclusion. Richardson, for instance, explicitly advocates for an open-ended conception of the moral community—one that does not rigidly define its boundaries by strictly determining which beings qualify as persons defined as entities that can wrong and be wronged (Richardson, 2018, p. 68).[445] More generally, the Specificatory theory requires that "only duty-holders, and not right-holders, need to be practically intelligent". While of course there is a moral desirability, for at least three main reasons,[446] on preferring the direct symmetry between two competent individuals, so that they can work things out together "when doing so is apt and feasible", any unavoidable asymmetries in the application of the Specificatory theory "will mostly arise from case involving right-holders, who are not morally competent individuals and who have no proxy or surrogate to speak for them. The idea that the Specificatory Theory is open to this possibility of our having duties to non-human animals, however, is an advantage, not a defect. […] Further work would be needed to settle which beings bear monadic rights." (Richardson, 2018, p. 113).

Bagnoli, instead, especially in some early papers, seems to be more restrictive regarding the concept of personhood or in general to the member of the moral community, by strictly connecting it with that of being an autonomous agents and also accountable (Bagnoli, 2007, p. 114).[447] Nonetheless, the main focus is on respect as a mode of recognition, "a practical relation

[445] Richardson admits to follow Michael Thompson's distinction (Thompson, 2004, pp. 351–355) who "illuminatingly contrasts persons, understood as beings who enter into such dyadic relations, with agents, who can be understood monadically. An agent acts on reasons, perhaps forming and pursuing a conception of the good, and perhaps even doing so rationally. While all persons are agents, not all agents are moral person." (Richardson, 2018, p. 67). This allows a broad range of views as to what beings might count as persons, even tough Richardson "do not defend any specific criterion of personhood or membership in the moral community" (Richardson, 2018, p. 68).

[446] See section 1.3.3 of the first chapter, and especially the footnote n.178.

[447] "Respect is a mode of recognition of others as persons, that is, as autonomous agents accountable for their actions. This recognition comes with reflexivity. Differently from non rational animals, we do not act directly moved by the urgency of desires. As rational beings, we question the authority of natural desires and we examine their legitimacy by considering whether they should be taken as reasons for action." (Bagnoli, 2007, p. 114).

that we entertain with ourselves and with others, rather than as a form of epistemic awareness." (Bagnoli, 2007, p. 117); the defining feature of a broader conception of personhood should, therefore, be found in the capacity for practical relationality and the recognition of interconnectedness. While Richardson thinks that emphasizing recognition above relational directed duties, as Bagnoli does, has indeed the disadvantage of being "quite restrictive as to the community membership" (Richardson, 2018, p. 64), Bagnoli concludes her book "Ethical Constructivism" precisely by looking forward to new developments of constructivism that could extend the question of membership in the moral community to animal, artificial and hybrid agents (Bagnoli, 2022a, p. 66).[448]

Whether this broadening of the moral community, along with the expansion of the concepts of personhood and dignity, will be pursued within the constructivist framework—and to what extent—remains an open question. The next few decades may bring crucial developments that will shape how constructivists can successfully accommodate these pressing challenges while maintaining the core principles of the framework. For the present scope, rather than offering a constructivist definition of personhood rights, we will remain focused on our original aim: providing a constructivist account of human rights. This can now be formulated as follows: Human Rights are reasonable demands that individuals specify toward one another in virtue of their shared moral status as bearers of dignity—a status that implies mutual respect. [449]

[448] "On the resultant picture, practical reasoning is not only productive but also self-transforming, and potentially socially empowering. A promising new direction of research exploits the metaphor of construction even further, uncovering the social normative dynamics of rational deliberation beyond any settled moral community. Can such a metaphor be usefully extended to account for novel normative communities, including various animal, artificial, and hybrid collectives? These are the challenges that lie ahead." (Bagnoli, 2022a, p. 66).

[449] We can, however, briefly sketch a possible constructivist definition of person rights by reformulating Richardson's account of the moral community (Richardson, 2018, pp. 63–89). In this view, Person Rights are reasonable demands specified by members of a moral community—that is, by individuals capable of wronging and being wronged, and thus bound by, and entitled to, a status of dignity that commands mutual respect.

PART II

HISTORICAL RECONSTRUCTION: TOWARD A CONSTRUCTIVIST CONCEPTION OF UNIVERSALITY FOR HUMAN RIGHTS

How do we know that human rights are not simply the invention of well-meaning minds shocked by the brutality of the Second World War and the horrors of totalitarian regimes? How do we know that they are what they promise to be – ultimate standards with cross-cultural, universal validity?

(Caranti, 2019a, p. 42)

That human rights is a uniquely Western concept is an exaggeration at best, a falsehood at worst.

(Baggini, 2018, p. 206)

In the first part of the project, we arrived at a constructivist understanding of human rights. They are reasonable demands that human beings specify among one another as possessors of a status that requires mutual respect: dignity. This understanding accounts for the procedural and relational nature of rights as outcomes of practical reasoning among moral agents (rather than as self-evident mind-independent facts) and explains the peculiar beneficiary elements of human rights, by clarifying in which sense the notion of dignity is a status-concept and not a value-concept. However, this definition still requires further development: to align with more established and widely recognized conceptions of human rights it must be supplemented by a more precise articulation of the additional features human rights are generally expected to possess.[450]

Despite the various definitions of human rights proposed by scholars across different schools of thought, one of the most widely agreed-upon characteristics is that human rights are universal and inalienable, meaning they cannot be granted or revoked by any authority since they apply to all human beings at all times and in all places, regardless of culture, nationality, or legal system. This characterization is indeed one of the central conceptual nodes of the Universal Declaration of Human Rights. The element of universality is often described as a legacy of the French Revolution, as a historical moment, and of jusnaturalism as a philosophical doctrine. The *Declaration of the Rights of Man and the Citizen* (1789) is generally taken to mark a turning point in European history, symbolizing a new way to view humanity based on a shared moral status (Corradetti, 2011). At the same time, the declaration originally maintained separated the "rights of man" and "rights of the citizen", a distinction that will be unified only in the *Universal Declaration of Human Rights* (1948). [451] This was the peak of natural law theory, but it also marked the beginning of its decline: the lack of a supranational authority "has limited the scope and application

[450] Ensuring such alignment is important not only for maintaining conceptual continuity with the normative practices through which human rights are recognized and enforced, but also for enabling critical engagement with those same practices. Only by situating theoretical reflection within existing human rights practices can we identify when those practices embody reasonable demands and when, instead, they fall short of the standards that a more coherent theoretical account might reveal.

[451] This duality reveals both the revolutionary promise and the limitation of Enlightenment rights discourse. On one hand, the declaration represented a profound shift toward recognizing human beings as bearers of a moral status—an early expression of what later became the idea of *human rights*. On the other hand, by maintaining the separation between "man" and "citizen," it effectively restricted the practical realization of universal rights to those who were politically enfranchised. Those excluded from citizenship—foreigners, women, slaves, and the stateless—remained outside the effective protection of these rights.

of the same Universal Declaration or, worse, has made its content obsolete in most of the cases" (Corradetti 2011, xv).[452]

For these reasons, Human rights are sometimes conceptualized as merely the political application of natural rights. Eugene Kamenka, for example, argued that the idea of human rights is clearly the successor of eighteenth-century European and American ideas of natural rights and thus grows out of Christianity, Stoicism, and Roman law; it differs, however, from earlier natural law doctrines in its emphasis on the individual and his or her values by elevating "the individual human person and his freedom and happiness to be the goal and end of all human association" (Kamenka & Tay, 1984, p. 6). In contrast, Samuel Moyn challenges the conventional view that traces the birth of human rights to the late eighteenth century, arguing instead that their current meaning and significance emerged more recently, around the 1970s. Moyn cautions historians and philosophers against focusing too much on earlier episodes in the rights tradition, as this could obscure what is novel about contemporary human rights (Moyn, 2012).

Rather than delve deeply into the historical evolution of human rights, which has been extensively explored,[453] this project will focus on a different inquiry, in order to reach a clearer conclusion to the main research question (in which sense a constructivist framework can contribute to the human rights debate); specifically, my aim in the II part of this work, is to consider whether a constructivist should regard the universality of human rights as grounded in the jusnaturalist tradition or instead seek to reframe it. To address this problem, In Chapter 4, I provide a comparative account of jusnaturalism and juspositivism through a historical and systematic examination of their most influential thinkers. In doing so, my aim is to show in what sense constructivism departs from both a jusnaturalist and juspositivist reading; I do not intend to suggest that the history of philosophy unfolds according to a linear or cumulative—let alone Hegelian—

[452] "When the Declaration of the Rights of Man and the Citizen of 1789 was proclaimed, Europe reached the apex of a historical turning point. For the first time it seemed that there was a meaningful way to address humanity as a whole, and that this was to be found within its common moral status. […] the separation between the rights of man and those of the citizen characterizing the French Declaration has found a "reunification" only with the Universal Declaration of Human Rights. Indeed, it was only in 1948 that the recognition of a universal right to take part in the government of one's country either by direct means or through the free choice of political representatives appears clearly (Art. 21.1). The Universal Declaration, therefore, is the highest level of expression of the natural law theory, setting a universal standard for rights to be respected without exceptions. The decline of a theory and a politics of human rights, though, began precisely from this apex" (Corradetti, 2011, p. xv).

[453] See especially Flynn (2011) for a detailed reconstruction.

progress. Rather, I wish to highlight how these conceptual oppositions have often been misunderstood or over-interpreted in the reception of classical authors, and how, in many of these thinkers, elements of both positions coexist. Once this aim is clarified, it should be clear that my approach is neither exegetical nor historiographical in the strict sense, but reconstructive: I seek to recover the arguments and normative stakes of these traditions, rather than merely to trace their chronological development or philological sources.[454] To this end, I draw both on the authors' own texts and on the interpretations offered by certain constructivist thinkers (Darwall, 2013b; Korsgaard, 1996b; O'Neill, 2015a).

This reconstruction will make it possible to challenge the common identification of human rights, natural rights, and universal rights, highlighting the risks of foundationalism inherent in certain jusnaturalism interpretations—risks that constructivism seeks to avoid without collapsing into purely formal or positivist notions of universality. The analysis will make it possible, in chapter 5, to address more directly the main question: "how can a constructivist recognize the universality of human rights?".

[454] Nonetheless, this is the part from which I benefited the most: having the opportunity to discuss and share drafts with various historians of philosophy with whom I was already in contact, as well as with the many new scholars I had the chance to meet at conferences during my doctoral studies. It was deeply rewarding to see that the initial intuition guiding my reconstruction—based on my general, textbook knowledge of the major philosophers, namely that many thinkers traditionally regarded as central to the natural law tradition also contained significant elements of legal positivism—was confirmed by several experts on each of these figures. I would like to thank all of them once again for their generosity and insight.

CHAPTER 4
CONSTRUCTIVISM AS AN ALTERNATIVE TO JUSNATURALISM AND JUSPOSITIVISM

The adjective "human" - unlike "natural" - does not suggest an ontological status independent of any and all human efforts, decisions, (re)cognition. It does not rule out such a status either. Rather, it avoids these metaphysical and metaethical issues by implying nothing about them one way or the other. The potential appeal of the select moral concerns is thereby further broadened in that these concerns are made accessible also to those who reject all variants of moral realism - who believe, for instance, that the special moral status of all human beings rests on nothing more than our own profound moral commitment and determination that human beings ought to have this status.

(Pogge, 1995, pp. 106–107).

4.1 A HISTORICAL REVISITATION OF A PERENNIAL DEBATE

4.1.1 The old idea of Natural Law

The concept of natural law seems to be as old as the very beginning of political reflection itself, since it was the simplest way to account for the seemingly paradoxical nature of an unjust law: If laws are the source of what is just, how can some law be unjust? This problem was already widespread among ancient Greek philosophers; both Plato and Aristotle were aware of how what was just by law was not always just by nature. However, two sophists, Antiphon and Hippias, deserve a notable mention, since in their system of thoughts the contrast between particular laws of the state and universal laws of nature emerges distinctly. Antiphon asserted precisely how "most

things that are right according to law are in opposition to nature." (Diels-Kranz, 87, B44, A 1-4),[455] because by nature the individual would pursue his personal benefit, which the law prevents. Corollary to this assumption by Antiphon was a thesis that was quite revolutionary for the time, namely, that since all men attend to natural necessities in the same way,[456] we are all naturally equal, barbarians included.[457]

This insight is embraced fervently by the Stoics, who place great emphasis on the rational aspect of human beings and assert that it is this rationality that renders them equal in nature. Rooted in a strictly monistic and teleological metaphysics, the Stoics tend to postulate a harmony between the law of the universe and the law of reason.[458] Everything is thus rationalized according to an order: individual things, even when they seem imperfect, have their perfection in themselves, embedded in the overall design. Man, as a part of the cosmos, lives according to the law of the cosmos and cannot rebel against his Nature. In this respect nature thus takes on the dual function of both referring to the nature of man and the nature of the universe, a nuance that will remain present in later uses of the concept, as well as one of the main reasons to the charge of vagueness for the term.

---

[455] The Diels–Kranz (DK) numbering is the standard system for referencing the works of the ancient Greek pre-Socratic philosophers, based on the collection of quotations and testimonies in "*Die Fragmente der Vorsokratiker*" (*The Fragments of the Pre-Socratics*) by Hermann Alexander Diels. The *Fragmente* was first published in 1903, subsequently revised and expanded three times by Diels, and finally revised in a fifth edition (1934–1937) by Walther Kranz. I am quoting the passages from the 2019 reprinted edition (Diels, 2019)

[456] "We all breathe the air with our mouths and nostrils, and we all eat with our hands." (Diels-Kranz, 87, B44, B 2).

[457] Assumption that Aristotle found reprehensible, but which was nevertheless a recurring element in the evolution of the debate on the subject. Reference to natural law concerned more specific particular norms that everyone was ready to accept, such as that killing another human being was wrong. More general principles, at least before the Stoics (who emphasized the natural equality of all rational beings), did not yet include a criterion of moral equality by nature; rather, the belief was that basic human beings were unequal by nature. For a historical-theoretical reconstruction on the issue of equality, see Gosepath (2021). For a critical reflection on the contemporary debate of the topic, see Phillips (2021).

[458] This is a key passage in the reconstruction of the natural law traditions we are providing. A constructivist perspective can be easily considered a jusnaturalist approach if by that we simply mean a universalistic conception of rights. However, regarding for example the specific stoic traditions, constructivism does not share any of the teleological elements. Unlike the Stoics, who perceive an inherent order in the universe alongside the faculty of reason, constructivism acknowledges no such order beyond the organizing and hypostatizing role of reason itself. The critical distinction lies in the fact that, for constructivists, universal laws are not discovered but rather constructed, a subtle nuance that is important to keep in mind as we delve into the works of the next authors like Aquinas.

The rationalistic turn of jusnaturalism by the Stoics reaches all the way to the systems of Christian thought, from the early Christian Augustine of Hippo, for whom natural law coincides with God's will, to High medieval period and Thomas Aquinas, who equates the natural law to the law of reason, in contrast with positive law.[459] As he wrote in Summa Theologica: "the participation of the eternal law by rational creatures is called the law of nature." (I, II, 91, 2).[460] Aquinas distinguished four types of law in total (*lex divina, lex aeterna, lex naturalis, lex humana*): natural law was understood as a reflection in human reason of the eternal law, which was the constitutive law of being, while the human law is the simple creation of contingent legislations (which can be just or unjust, based on how they reflect the natural law). Natural law can be discovered by reason alone and applies to all people, while divine law can be discovered only through God's special revelation and applies only to those to whom it is revealed and whom God specifically indicates are to be bound.

In this sense, Aquinas' conception of natural law is twofold and depends on the point of view of the observer of that law: when one focuses on the role of God as the giver of natural law, natural law is only one aspect of divine providence (and thus the theory of natural law is from this perspective only one part among others of the theory of divine providence); when one focuses on the role of man as the recipient of natural law, natural law constitutes the principles of practical rationality, those principles by which human action is to be judged as reasonable or unreasonable (and thus the theory of natural law is from this perspective only the preeminent part of the theory of practical rationality). The idea that natural law constitutes the fundamental principles of practical rationality implies, for Aquinas, both that the precepts of natural law are universally binding by nature and that the precepts of natural law are universally knowable by nature (Murphy, 2019).

[459] Although both concepts were used by Greeks, and especially by the Stoics, and it was Abelard who introduced the term 'positive law', it was thanks to Aquinas that the term gained a wide philosophical circulation (Finnis, 2020). Aquinas marks a turning point: "the attempt to understand morality in the legalistic terms of a natural law is ancient but is now mostly associated with the formulation given it by Thomas Aquinas in the late thirteenth century. Earlier natural law is commonly seen as leading up to Aquinas's paradigmatic version, whereas later natural law is understood as deriving from it." (Haakonssen, 1996, p. 15).

[460] I–II, 91, 2 refers to *Prima Secundae*, Question 91, Article 2, following the traditional division of the Summa Theologiae

In light of all this, it seems that jusnaturalism and juspositivism are already present and intertwined in this conception, in contrast to the classical interpretation where Thomas Aquinas presents natural law as a rigid, rational order imparted by God to the cosmos, passively received by humans. In this traditional view, foundationalist and formalist elements are seen as opposed and mutually exclusive. Accordingly, the authority of laws is either the product of human reasoning or an objective normative element discovered in the nature of things. Aquinas is often interpreted as aligning with the latter, suggesting that the rightness of rights is not a result of reasoning but rather an objective normative truth inherent in nature. However, for Aquinas, there is no inherent contradiction between these two dimensions: natural law is both the divine law given by God and the framework through which we reason and come to understand this very law.

The apparent circularity need not be a vicious one if our understanding of law arises not from mere discovery but through a process of reflexive reasoning. This seems to be Aquinas's approach when he speaks of "progressive illumination" and argues that faith builds upon, rather than contradicts, what can be known through reason. While faith is indispensable for apprehending certain theological truths about God—such as the Trinity or the Incarnation—it is not required for grasping the natural and moral order.[461] This reflexive element, which will later be rendered in more procedural and secular terms by subsequent thinkers, offers one possible way to move beyond the traditional dichotomy between jusnaturalism and juspositivism.

4.1.2 A Modern Conceptualization of Jusnaturalism

The transition from the Middle Ages to Modernity brought about significant changes across various spheres, including the legal domain. These shifts were marked by the decline of religious unity and the diminishing influence of Roman law, which gave way to the rise of *ius patrium* (local or customary law) and the rethinking of legal principles and their application. Law began to be viewed less as an immutable tradition handed down by authorities and more as something subject to limited interpretation by judges, who were increasingly expected to apply codified, written laws.

[461] In "*Summa Contra Gentiles*", *Book I, Chapter 3* he wrote that: "The truths that we confess concerning God fall under two modes. Some things true of God are beyond all the competence of human reason, as that God is Three and One. Other things there are to which even human reason can attain, as the existence and unity of God, which philosophers have proved to a demonstration under the guidance of the light of natural reason"

This shift reflected a move toward formal legal structures where judicial discretion was curtailed in favour of adherence to written statutes. In this context, the adjective 'natural' took on a dual polemical significance. First, it opposed the growing prevalence of codified law, which often represented a layered accumulation of privileges and social inequalities; natural law, by contrast, was seen as a rational, universal system that denied such privileges, asserting equality and fairness as self-evident truths inherent in human nature. Second, the concept of natural law stood in opposition to the absolutist state, which had begun to centralizes power and restrict individual freedoms.

Arguably, Hugo Grotius is the first of the modern thinkers to make a major contribution to the desacralization of legal orders, describing nature as belonging to the domain of human reason.[462] For Grotius there are three institutions capable of producing law: the international community, states and families (in order of decreasing authority). The universal laws of nature serve as the foundation of international law, holding valid for all states regardless of their positive legal systems. Grotius is also the first scholar to use the word "right" to refer to a moral quality inherent in person (a quality that gave them certain privileges) and he also introduced the concept of sociability as "the fountain of right" (Darwall, 2013a, p. 186). He believed that a fundamental aspect of the universal laws of nature—the natural right of all individuals and groups to self-

[462] In considering Grotius as the first modernist, I am drawing on Stephen Darwall (2013a) and Jerome Schneewind (2007), but even Pufendorf remarked how Grotius should be considered the true founder of the modern school of jusnaturalism. Nonetheless, between St. Thomas Aquinas and Grotius, other important figures should be mentioned. Francisco Suarez, for example, as noted by Darwall, brilliantly observed that deontic concepts related to moral law and obligation involve specific elements of accountability not inherent in the concept of a normative reason. Saying that moral reasons advise an action is different from saying an action is morally obligatory. Similarly, saying an action is morally permissible is different from saying it is not obligatory to refrain from it. While Suarez agreed with much of Aquinas's natural law theory, he believed Aquinas missed how natural law imposes obligations through authoritative demands that people are accountable for following: moral norms are god's will as addressed to us and our rational will (Darwall, 2013b, p. 163). This focus on the second-personal structure of obligation and the relationship between accountability and authority mark a crucial step toward the constructivist understanding of laws and norms. By contrast, another important figure at the beginning of the modern period is Niccolò Machiavelli, whose work explicitly endorses the separation between morality and law. As a matter of fact, "The Prince" focuses on the strategies and tactics that rulers use to acquire and maintain power, often disregarding traditional moral norms. Importantly, the book was meant to be a guide for how princes ought to behave, not just a description of how they actually behave. This is what made it scandalous: everyone knew that Renaissance princes acted immorally, but only Machiavelli dared to argue that they were justified in doing so (J. P. MacCormick, 2018, p. 3). This aligns more closely with an early form of what is now referred to as political realism, which is distinct from juspositivism, but is equally problematic for constructivist scholars.

preservation, a right universally recognized—could serve as the basis for resolving both conflicts between nations and disputes between private individuals in peacetime.

As Darwall underscores, what is groundbreaking in Grotius is the notion that morality (natural law) establishes "obligations" whose binding force cannot be reduced to mere reasons or recommendations of any kind (Darwall, 2013a, p. 158).[463] Grotius himself explicitly points out that "ancient" conceptions of right and law lack any conceptual connection to obligation and legitimate demand when he "contrast the Aristotelian conception of distributive justice that "ancients" and even some "moderns" include under the concept of right with right "properly speaking", since the latter includes doing for others "what in Strictness the may demand" (I, 88-9)" (Darwall, 2013a, p. 176).[464] Thus, the problem of the thinkers preceding Grotius is, for Darwall. that they believe that natural law and agent own good coincide, either in an Aristotelian or catholic sense or both, whereas Grotius hold that natural law does not simply add an obligation,[465] but a normative, internal, reason.[466] And yet, this internal reasons is expressed in a "moral community as mutual accountability" (Darwall, 2013a, p. 183) where individuals are morally obligated to hold each other accountable for their actions. A person's natural right to life thus does not simply creates a corresponding obligation for others not to harm them, but issue a legitimate claim on other members of the moral community.

Interpreting Grotius as "holding that the law and right of nature are grounded in the capacity of rational persons to recognize their common competence and authority to make reasoned claims and demands against one another and to live with one another on terms that respect this common standing" (Darwall, 2013a, pp. 187–188) positions him as the true precursor to

---

[463] "Moreover, central to morality as Grotius understands it, are basic individual "perfect rights" all persons have, *claim rights,* as we would now call them, which entail the standing or "faculty of demanding" rightful treatment from others and others' consequent obligation to provide it (Grotius 2005: I, 139)." (Darwall, 2013a, p. 158).

[464] For Darwall, something can be a natural law of right, in the proper modern conception, only if it entails *normative* reasons for agents to act (Darwall, 2013a, p. 173).

[465] For example, against Suarez he holds "that natural law does not simply add an obligation (of obedience to God) to what is already morally right. What grounds the laws is not a superior's will, but our common rational and sociable nature conceived already in "juridical or quasi-jural" terms" (Darwall, 2013a, p. 171).

[466] Darwall recall the contemporary division of the debate by distinguishing how thinker that precedes Grotius, such as Suarez, can be considered externalist, since reason for compliance must come from the agent's own good, while Grotius can be viewed as an internalist, since what grounds "natural law and right as genuinely obligating must simultaneously supply normative reasons for compliance" (Darwall, 2013a, p. 177).

modern though. By linking morality to the concept of legitimate demands and distinguishing its normative authority from that of reasons which can only recommend or counsel conduct, albeit strongly, without yet compelling it, Grotius introduced into the modern era the problem now known as the "authority of morality" — namely, how to understand and justify morality's unique normative force (Darwall, 2013a, p. 158). Nonetheless, while Grotius famously claimed that natural law would still hold even if God did not exist (*etiamsi daremus non esse Deum)* he did not fully address the implications of the mechanistic and anti-teleological worldview that came with the Scientific Revolution. This is why Korsgaard contends that the credit often given to Grotius should instead be attributed to Thomas Hobbes and Samuel von Pufendorf, since "they were the first to identify clearly the special challenge which the Modern Scientific World View presents to ethics, and to construct ethical theories in the face of that challenge" (Korsgaard, 1996b, p. 21).

Indeed, Thomas Hobbes fundamentally reoriented the basis of moral and political obligation from a natural law rooted in universal command to one centred on the natural desire for self-preservation by humans. In contrast to earlier thinkers like Grotius, who emphasized human sociability as a natural basis for morality and law, Hobbes argued that humans are not naturally sociable: in their natural state—the "state of nature"—individuals are driven by self-interest, leading to a condition of constant competition where each person has a right to everything. This results in a perpetual state of conflict or "war of all against all" (Hobbes, 1651, Chapter I, Xiii). For Hobbes, the fundamental law of nature is that individuals should seek peace, but only to the extent that peace is compatible with self-defence. This understanding of natural law is not based on any inherent moral qualities but is rather a practical principle grounded in survival: the first natural right in Hobbes' framework is the *conatus sese conservandi*—the drive for self-preservation. To escape the violence and instability of the state of nature, individuals collectively agree to a social contract, surrendering their natural rights to a sovereign authority in exchange for protection and peace. Moral obligations, in this view, are not independent of human will but are products of the social contract, which exists to ensure self-preservation. Right and wrong are determined by the laws established by the sovereign, whose authority stems from the need to maintain social order.

It is notable that Hobbes is perhaps the first thinker to spark significant philological debate regarding his stance on the jusnaturalist-juspositivism debate. His famous assertion that the first

law of nature is to seek peace and defend oneself highlights the centrality of natural law as a fundamental principle derived from human nature. However, his emphasis on the social contract as a mechanism for individuals to escape pre-institutional chaos, along with the importance he places on authority and governance in creating and enforcing laws, has led some scholars—such as Hart and Raz—to associate him with legal positivism.[467] This ambiguity is unproblematic when we recognize that most authors incorporate both universal and procedural elements into their theories, even if they were not already able to reunify the two. And it becomes even less troubling once we acknowledge— as constructivists do— that some form of moral grounding is inescapable also for juspositivists.[468]

Building on the work of both Grotius and Hobbes, Samuel von Pufendorf synthesized their ideas, aiming to address their limitations and reconcile human sociability with the emerging scientific worldview. Like Grotius, Pufendorf conceptualized natural law as applicable to both individuals and political communities, placing sociability at the forefront of his framework, as a trait inscribed in human nature, discernible through natural reason. However, he also incorporated Hobbes' insight that this sociability has a practical dimension: human beings are not sociable because of an inherent moral good, but because living in society is necessary for survival and well-being (Pufendorf, 1717). This dual conception of sociability as both a natural inclination and a practical necessity represents Pufendorf's solution to the apparent dichotomy we have just discussed: human beings are naturally inclined to live in society, but this inclination is driven by the pragmatic need for security and cooperation, rather than by any innate moral virtue (Schneewind, 2007, pp. 170–172).[469] Thus, for Pufendorf, sociability serves as a practical task that

[467] Against this reading, see especially James Boyle (1987).

[468] Hobbes obscured this inescapability by reducing morality and the "laws of nature" to purely prudential reasoning concerned exclusively with what serves one's self-interest (Hobbes, 1651, Chapter I, xv). Such reasoning, he argues, leads us to pursue civil peace and to endorse the political arrangements required to secure it—above all, the establishment of a sovereign with coercive authority. But, as Charles Larmore rightly observes, "What Hobbes overlooked are the moral assumptions underpinning this line of argument. For instance, it embodies an instrumentalist conception of morality that rejects various moral concerns, such as being impartial between the powerful and the weak and valuing right action for its own sake, that many would think essential" (Larmore, 2020, pp. 42–43, fn 16). (Larmore, 1996, 2020)

[469] According to Pufendorf, the actions of human beings, like every other form of physical motion, are in themselves morally indifferent: "values are not found in the world of nature at all".(Korsgaard, 1996b, p. 21). Instead, Pufendorf says, intelligent beings must impose moral values on nature: "as the original way of producing physical entities is creation, so the way in which moral entities are produced can scarcely be better expressed than by the word

must be cultivated for the sake of collective survival, giving it a normative dimension even if it is not grounded in a moral imperative, but precisely because of its practical indispensability.[470]

In Pufendorf's system, the laws of nature are both practical guidelines for living together in peace and principles that arise from the necessity of cooperation in a world shaped by human needs and interests. In this way, Pufendorf reconciles Hobbes' focus on self-preservation with Grotius' emphasis on the social nature of humans, showing that sociability is essential for both individual survival and the functioning of political communities. By doing so, he bridges the gap between a purely survival-driven account of human relations and one that recognizes the inherent necessity of social life, providing a more comprehensive foundation for the development of moral and political obligations. Even though for Darwall, this ends up in a "fascinating, but unstable, balance of elements from Grotius and Hobbes",[471] it is nonetheless a "fascinating" mix of "reciprocal recognition and accountability" (Darwall, 2013b, p. 221).[472] In this sense, Pufendorf's

imposition. For they do not arise out of the intrinsic nature of the physical properties of things, but they are superadded, at the will of intelligent entities, to things already existent and physically complete, and to their natural effects." (Schneewind, 2007, p. 171).

[470] "As Pufendorf understands inherently "social" attitudes of kindness, peace, and loves, these already involve a kind of mutual recognition and respect" (Darwall, 2013b, p. 201).

[471] "Like Grotius, Pufendorf emphasizes sociability which he ties explicitly to the recognition of equal dignity and mutual rights, something that is only implicit in Grotius. But whereas sociability is foundational for Grotius, it is derivative for Pufendorf, and in a way that I have argued makes Pufendorf's view unstable. On the other end, a connection between morality and the "natural" power to impose sanctions, which is so important for Hobbes in creating a robust realm of obligations in *foro externo*, is no less essential for Pufendorf. But here again, this introduces instability into Pufendorf's theory, since he is at pains to distinguish his moral psychology and view of the natural moral condition from Hobbe's" (Darwall, 2013b, p. 221). While Darwall's reading may seem somewhat critical of Pufendorf in favor of Grotius, he convincingly highlights the important distinctions Pufendorf draws in relation to both Grotius and Hobbes. For instance, while both Grotius and Pufendorf view sociability as fundamentally involving relationships "conceived in terms of mutual recognition of individuals' respective authorities or normative "powers", that is, in terms of their capacities to have obligations to and rights against one another" (Darwall, 2013b, p. 192), Pufendorf diverges from Grotius conception in a significant way, arguing that without God's authoritative will "neither natural law in general nor right and ("directed" or "bipolar") obligations between human beings more specifically, could possibly exist." (Darwall, 2013b, p. 192). Similarly, Pufendorf rejected Hobbes's notion of a natural right to all things, where individuals in the state of nature could do anything necessary to preserve their lives, including taking resources or harming others. This is not a "real right" since it entails "no correlative moral effect in the rest of mankind"(Darwall, 2013b, p. 205). In contrast, Pufendorf believed that natural law imposed moral duties on individuals, such as living peacefully with others: this meant people were bound by obligations to respect the rights of others, even in the absence of a formal government and that liberty was not unlimited in the state of nature. Thus, the fundamental, and often overlooked, shift is from Hobbes's focus on self-interest to Pufendorf's conception of moral power. "Exercising moral power creates new right and obligation […] We can of course say that a general obligation to keep promises or comply with authoritative orders already exists, but any such general obligation will itself depend on the normative powers to create voluntarily the more specific bipolar obligations in which promise and authority themselves respectively consist" (Darwall, 2013b, pp. 194–195).

account is as a crucial bridge between early modern philosopher and contemporary political scholars.

The same year that Pufendorf was born was the same as John Locke[473], who is usually considered the most important jusnaturalist of the modern times.[474] Locke argued in detail, mainly in the *Two Treatises of Government*, that certain rights self-evidently pertain to individuals as human beings (because they existed in the 'state of nature' before humankind entered civil society;) that chief among them are the rights to life, liberty (freedom from arbitrary rule,) and property. Like Pufendorf, Locke argues, as part of his criticism of innate ideas, that moral knowledge is not implanted in the mind by nature, but constructed by rational agents: moral ideas are "mixed modes", namely complex ideas deliberately put together from simpler ones to describe and evaluate human actions (Locke, 1690, Chapter II.xxxi.3; III.v.3; III.v.7; III.xi.16).[475] Moral

[472] "Pufendorf was far from the first thinker to hold some version of a divine command theory of morality. But he may have been the first to try to work out what such a view must look like if it is to take seriously the conceptual links between authority, recognition, and accountability, as well as the psychological necessary for these to be realized in the moral life." (Darwall, 2013b, pp. 189–190).

[473] Again, in this brief historical overview I am concentrating on the key authors who are most pertinent to the specific aspects of the topic I wish to explore. However, it is still worth to at least mention some notable figures of this historical period that we unfortunately cannot examine in detail. For instance, in the period between Grotius and Hobbes, John Selden (1584-1654) stands out; between Hobbes and Pufendorf/Locke, several notable figures such as Nathaniel Culverwell (1619-1651), Richard Cumberland (1631-1718), and soon thereafter Wilhelm von Leibniz (1646-1716) warrant mention. For a more detailed examination of these and other modern thinkers on the subject of natural law, one can refer to Tuck (1982) and Haakonssen (1996).

[474] As we will discuss in the following pages, particularly in the next two notes, most scholars are skeptical about this claim. Locke is often labeled a textbook natural law theorist, primarily in contrast to Hobbes (Parkin, 2007) since Locke does not deny the existence of good and evil in the state of nature as Hobbes does. However, this raises a significant issue, as both Locke and Pufendorf hold a voluntaristic view of law—that every law must be an expression of a legislator's will. If there is no true legislator in the state of nature to establish laws, how can natural law exist? The solution both Locke and Pufendorf offer is that natural law in the state of nature can indeed reflect a legislator's will because it is established by God in accordance with human nature. This divine law is not purely voluntaristic (it is not arbitrary; it could not be otherwise) but aligns with a broader natural order and with the inherent characteristics of human beings. Locke's resolution seems thus to blends both perspectives, which he also acknowledges as a "strange doctrine" (Locke, 1690, p. II.7 § 9); the duties of natural law are accessible through reason, and to uphold natural law without a legislator, Locke grants individuals the right to punish transgressions themselves. "The legislator is necessary to make obligation possible, that is, to make morality normative." (Korsgaard, 1996b, p. 27).

[475] In much of his philosophy relating to morals Locke is working out a system similar to Pufendorf 's, and, in view of Locke's stated admiration for the German thinker, he is likely to have been doing so with some degree of self-consciousness (Haakonssen, 1996, p. 52). Their meta-ethical convictions significantly shaped how Pufendorf and Locke contributed to the evolution of natural law and natural rights philosophy, shifting the discourse away from divine or inherent power structures and emphasizing human reason as the foundation for moral and political principles. However, while Pufendorf grounded natural law in the inherent social nature of human beings, Locke centered it on individual rights that exist prior to, and independently of, societal structures. This distinction highlights their different approaches to the relationship between human nature, reason, and moral obligation.

ideas and moral rights are, in this sense, tools that help us make sense of our practical and social experience, but their universality does not seem to come from their normative authority per-se,[476] but through a practical process of human reasoning: reason allows humans to recognize their equality and independence, and therefore, that no one ought to harm another in their life, health, liberty, or possessions.[477]

This is why according to Locke, upon entering civil society humankind surrendered to the state only the rights to enforce these natural rights, not the rights themselves. If a government fails to protect these rights or acts against the common good, citizens have the right to revolt.[478] Locke hence acknowledges that the enforcement of human institutions in creating and maintaining laws was a necessary step. This is exemplified by his treatment of the right to private property: while property rights are derived from natural law and based on the labour one invests in nature, the regulation and protection of property require positive laws enacted by a government. The interplay

---

[476] The continuity of ideas between "*An Essay Concerning Human Understanding*" and the "*Two Treatises of Government*" is a major point of debate among Locke scholars. In the *Two Treatises*, Locke appears to present a natural law theory (there exists an immutable and eternal law of nature, and natural rights derivable through reason); in the *Essay*, however, he argues that moral law depends on a higher authority who enacts the law and possesses the right and means to enforce it through punishment and reward, asserting that the goodness or evil of actions depends on their alignment with a given law (Locke, 1690, p. II.xxviii.5-6). Regardless of how one interprets Locke to make his views more internally consistent, his position ultimately remains a blend of the two perspectives—similar to that of Pufendorf.

[477] Even a thinker like Locke can thus be reconceptualized as a moral constructivist? More modestly, we might say that contemporary metaethical categories like constructivism and realism don't fully capture the complexity of his thought. Locke's emphasis on rational deliberation and human agency suggests elements of constructivism, but this doesn't obscure the fact that other aspects of his philosophy remain aligned with moral realism, especially where moral facts—such as the wrongness of harming others—appear grounded in human nature and in the divine law that generates it. This foundational appeal to a theologically anchored moral order would be incompatible with a constructivist approach that seeks to accommodate reasonable pluralism in modern, diverse societies. In fact, unlike Locke's Britain, which was religiously homogeneous enough for him to assert that "those are not at all to be tolerated who deny the being of God" (Locke, 1689, p. 313), contemporary liberal societies are marked by deep religious and moral diversity. They extend equal legal and political recognition even to those who reject theistic frameworks altogether. While Locke could claim that "the taking away of God, though but even in thought, dissolves all" (Locke, 1689, p. 313), constructivist and liberal theorists today must seek more inclusive foundations for political legitimacy and moral obligation.

[478] Flikschuh (2008, p. 379) underlines how there is "an arresting shift in Locke from an individual right to self- (and other) defense in the pre-civil condition to the people's right to revolution in the civil society. It is not clear how Locke performs this shift." We will return to Flikschuh's reading more extensively in the next subsection on Kant, but it is important to highlight here a crucial contrast: unlike Locke, for Kant civil society does not precede the formation of the state—it becomes possible only through it. Prior to the establishment of the state, there exists merely a multitude (Flikschuh, 2008, p. 382).

between pre-existing moral rights and the practical need for a legal framework to manage and protect these rights was seen by Locke as a necessary synthesis.

Already in Locke, therefore, the force of universal rights seems to be legitimized no longer by their ontological status as a realist basis nor by arbitrary creation ex nihilo, but by the combination of these elements, that is, a formal procedure of construction that has constitutive foundations in the very process of rational deliberation. This is also one of the often-overlooked central points of Immanuel Kant's philosophy, who is often considered the principal, manifesto philosopher of jusnaturalism (Passerin d'Entrèves, 1970).[479] Yet there are also authors who speak of Kant as a clear jus-positivist (Brudner, 2011; Waldron, 1996)[480], especially when he states that one cannot speak of rights unless there is a state that recognizes them, a statement that has been read by some (Nussbaum 1997, 273) as subscribing to the jus-positivist thesis at the expense of the jusnaturalist one. How can such ambivalence be explained?

It is however essential first to briefly mention some Enlightenment thinkers, from the period between Locke and Kant, to understand more clearly Kant's intellectual background. This includes the Scottish Enlightenment, exemplified by David Hume (1711-1776) and Adam Smith (1723-1790), as well as the French Enlightenment, represented by Voltaire (1694-1778) and Jean-Jacques Rousseau (1712-1778). Scottish thinkers especially emphasized the importance of sociability, common sense and emotions alongside reason.[481] In particular, David Hume's

[479] "Kant was by far the most incisive exponent of natural law theory in the modern era", and as such he was also "the most consistent and persuasive critic" of legal positivism (Passerin d'Entrèves, 1970, p. 110).

[480] "If, nevertheless, there are reasons for thinking that society needs just one view on some particular matter to which all its members are to defer, then there has got to be a way of identifying a community view and grounds for one's allegiance to it that are not predicated on any judgment one would have to make concerning the view's moral rectitude. That is the positivist position, and Kant's contribution is to have linked that position conclusively - in the way we have traced - to the idea of a set of property rights, secured and assured as a system" (Waldron 1996, 33).

[481] "For these thinkers the moral life was - eschatology apart - an open-ended series of contractual adaptations among individuals. […] Hume and Adam Smith became part-heirs to this tradition by transforming contractarian negotiation into a complex socio-political development that had to be understood historically. In general, however, natural law theory gave a very different idea of the moral condition of humanity. According to most natural lawyers in the seventeenth and eighteenth centuries, moral agency consisted in being subject to natural law and carrying out the duties imposed by such law, whereas rights were derivative, being mere means to the fulfilment of duties. Each person had a right to the acts he or she was under a duty to do and a right to what others had a duty to render. Similarly, contracts were means for implementing the performance of the duties imposed by natural law. Within this conceptual framework an intense battle was fought during the last three decades of the seventeenth and the early years of the eighteenth century. At issue was the old problem of whether natural law had moral force for humanity solely because it was God's will or whether in addition it had independent moral authority with us. Few disputed that natural law existed because of God's will; the question was rather whether or not there were moral values shared by God and

empiricism and skepticism about the rationalist foundations of knowledge notably prompted Kant to critically examine the limits of human reason. Similarly, Adam Smith's emphasis on education influenced Kant's concept of *Bildung* (self-cultivation or formation);[482] for Kant, however, moral education was not about teaching specific virtues but rather about fostering individuals' capacity to reason about what is morally right.

French Enlightenment thinkers like Voltaire and Rousseau also underscored the role of education in developing a person's capacity for autonomous moral reasoning, emphasizing individual rights and autonomy. Kant expanded on these ideas by making autonomy central to his conception of moral law, arguing that true morality arises when individuals legislate moral laws for themselves, independent of external authority. By synthesizing all these influences, Kant developed a peculiar conception of the jusnaturalist framework that emphasized both the social and a-priori dimensions of moral obligations.

4.1.3 Kant as the Bridge Between Jusnaturalism and Juspositivism

To determine whether Kant can be accurately considered a jusnaturalist, one way of proceeding is to examine how and to what extent Kant engages with the vocabulary of natural law. In the *Groundwork of the Metaphysics of Morals* (1785), the first work of moral philosophy published by Kant in his critical period, he articulates the supreme principle of morality by calling it not only a universal law, but also a law of nature (G 4:421). In fact, Kant introduces many formulas of the categorical imperative, the most notable being "act only according to that maxim through which you can at the same time will it to become a universal law [ein allgemeines Gesetz]"

humanity which entailed the moral obligations of natural law independently of our regard for God's willing this to be so. This was closely associated with the question of what powers of moral discernment human beings possessed, and that was in effect a question of the impact of original sin. The mainstream of moral philosophy in the Scottish Enlightenment was formed by dealing with this issue. The question of the adequacy of humanity's moral powers was answered empirically in the well-known theories of 'moral sense' and 'common sense'". (Haakonssen, 1996, pp. 5–6)

[482] "The law of nature points out what the highest good is, namely the pleasures of a possibly eternal life, but the law does not constrain the will to do the good and thus terminate the individual's liberty. Liberty consists in having compliance with the law as a rational option. This leaves scope for error and thus raises the need not only for sanctions but also for education"(Haakonssen, 1996, p. 56).

(G 4:421); but also immediately afterwards he restate it as: "act as if the maxim of your action should by your will become a universal law of nature [allgemeines Naturgesetz]" (G 4:421).

Kant thus seems to continue this tendency of conflating the concept of universal law with that of natural law, an overlap that is then contrasted en bloc with positive law. Indeed, in the *Metaphysics of Morals* (1797), he also explicitly distinguishes between natural rights, based solely on a priori principles, and positive (statutory) rights derived from the will of a legislator (MS 6:237). There is here, however, an important nuance that is easily overlooked: for Kant, a priori principles are *formal* rather than *substantive*, and in this sense, they are "empty" with respect to the particular domains they legislate. Unlike some jusnaturalists conception of natural law based on human nature, teleology, or natural purposes, Kant explicitly rejects such foundations as heteronomous, namely dependent on contingent and empirical ends, rather than to the autonomy of pure practical reason. Kant's concept of natural law needs therefore to be interpreted not in a foundationalist sense, as teleologically determining the will, but in a descriptive sense, as shaping and clarifying action.[483]

What is crucial, therefore, is to determine the descriptive sense of natural law and the obligations that arise from it. Kant is very clear that a duty (*plifchten*) is the necessity of an action done out of respect for the moral law. This respect for the moral law is a form of self-legislation, where individuals recognize and follow laws that they themselves are bound to, independent of external influences, incentives, or consequences. In *The metaphysic of Morals*, which is divided precisely into two main parts, the Doctrine of Right (*Rechtslehre*) and the Doctrine of Virtue (Tugendlehre), Kant distinguishes between two types of duties: duty of rigjt (*Rechtspflichten*) and duty of virtue (*Tugendpflichten*). What essentially distinguishes these two duties is the fact that

[483] Many scholars remark that this emphasis on the notion of law is due to Kant's Newtonian conception of reality: since every form of causality in nature needs a law, the moral law is the descriptive law of freedom. Furthermore, Schneewind suggests that Kant's focus on law can also be viewed as a response to prominent thinkers like Grotius and Pufendorf (Schneewind, 1998). For Kant, the foremost normative reality is indeed a law, but it is critical to note that this law does not originate from an external source – such as God, the state, or other individuals. Adherence to such an externally imposed command would constitute heteronomy. Instead, the law that Kant refers to is one that emanates directly from reason itself. This intrinsic law is self-governed, highlighting Kant's unique contribution to the understanding of moral and ethical principles.

external compulsion to a juridical duty is morally possible, whereas a duty of virtue is based on free self-constraint out of moral commitment (MS 6:383).[484]

Clearly, this distinction can be interpreted in a twofold sense.[485] A juspositivist might underline how it confirms that these are entirely separate domains, and that legal validity does not depend on their moral content (Ryan, 1986, p. 79).[486] This separation seems indeed problematic for jusnaturalists, especially because it is difficult for them to accommodate Kant's assertion that some duties are morally binding but are not enforceable through external authority (Nussbaum 1997, 273) and especially that legitimate enforcement of pre-civil right claims is possible only in the civil condition and this is why, contrary to Locke, there is no right to revolution or to resist the state (TP 8:300). Nonetheless, jusnaturalists can also interpret many passages, and even the whole structure of the Metaphysic of Morals, to show that even if ethics and legality diverge in important ways, they are still somehow importantly connected and the fact that they are both parts of the Metaphysics of Morals shows that (Mulholland, 1990). For juspositivist that endorse the separation

[484] As Fiorella Tomassini (2024) perceptively demonstrates, Kant's approach represents an innovative synthesis that critically engages with two dominant natural law theories of his era, Achenwall's coercion-based model of right and Baumgarten's duty-oriented system. From Achenwall, he adopts the crucial distinction between right and ethics, along with the conceptual linkage of juridical authority to coercive power. Simultaneously, he incorporates Baumgarten's systematic vision, which situates rational right within a comprehensive framework of duties. This dual appropriation allows Kant to forge a novel theoretical position that transcends the limitations of both predecessors while preserving their most valuable insights.

[485] As Alan Wood succinctly captures: "Kant's Metaphysics of Morals divides 'morals' (Sitten) into two parts: right (Recht) and ethics (Ethik). But Kant does not make it clear how right and ethics relate to each other, or how their treatment forms a single whole." (Kisilevsky & Stone, 2019, p. 143). This ambiguity in inherited regarding the relationship between natural and positive law. See the next note.

[486] "Thus we find Kant both asserting as vigorously as any legal positivist that men only have property in external things when a legal order gives them that property and provides remedies for its loss, and asserting that we have to assume a 'natural right' to appropriate unowned things and make them our property in a state of nature" (Ryan 1986, 79–80). Ripstein also recognize this ambivalent opacity, but, as I will argues, he does so precisely to overcome both the outdated categories of jusnaturalism and juspositivism. As Ripstein puts it: "Kant lies outside the primary canon in the history of political philosophy, in part because his writings on the topic are frequently opaque, even by the standards of his other writings. That opacity often leads readers to take him to hold positions that are not importantly distinct from those of others in the natural law or social contract traditions. So he is sometimes compared, or assimilated, to Aquinas, Grotius, or even the elusive Pufendorf. More frequently, his work is lumped together with that of Rousseau, or taken to occupy some sort of middle ground surrounded by the trio of Hobbes, Locke, and Hegel. Yet Kant's political writings, particularly in the Doctrine of Right, resist assimilation to either the natural law or social contract traditions. For both of those traditions are rooted in a pair of ideas that Kant rejects." (Ripstein, 2015, p. 42).

thesis, this is difficult to handle, because it challenges the idea that legal norms can be valid without reference to moral considerations (Kisilevsky & Stone, 2019, p. 172).[487]

To summarize this broad and obscure debate, we can say that jusnaturalists reading of Kant tend to focus mostly on the equivalence between universal and natural law as a moral foundation (G 4:422) and on the *relation* of natural and positive rights (MS 6:237) while juspositivist reading will focus on the *separation* of natural and positive rights (MS 6:237) and on law's coercive authority and the impossibility of a right to resist the state (TP 8:300). In response to this tension, some scholars try to overcome this tension between the jusnaturalists and juspositivists reading, by reaffirming the unity of the law and ethics through their common foundation not in a natural theory sense, but in practical reason. I will focus briefly here especially on the reading of Arthur Ripstein (sometimes considered a jusnaturalist reading)[488] and Katrin Flikschuh (sometimes considered a juspositivist reading) and in the end I will also show that the constructivist reading adheres with this general alternative interpretation, by examining a specific paper of O'Neill on the role of the social contract in Kant.

In his monumental work *Force and Freedom* (2009) Arthur Ripstein asks how coercion can be compatible with each human being's right to be their own master in a Kantian framework. According to Ripstein, Kant's political philosophy—particularly as developed in the *Doctrine of Right*—does not rest on the application of general moral principles to factual circumstances, nor

[487] "Should this package be characterised as positivism? It can be; the present point isn't to legislate the use of a word. What matters is just that two points be kept in mind. First, all legal theories acknowledge a role for positive law—a role for legal norms identifiable merely on the basis of their sources. What characterises traditional 'natural law' theories is that positivity does not exhaust the concept of law—law has also a critical, moral part—and by this standard, Kant continues the natural law tradition: he endeavours to situate the law's positive aspect within morality, to exhibit it as a moral requirement. Second, a contemporary positivist is typically arguing with someone who claims that independent or self-standing moral truths can or must determine the content of the law. But Kant's version of the continuity of law and morality isn't a matter of the law taking marching orders from self-standing moral principles. As I have emphasized, the Principle of Right is proprietary to law and dependent on law for authoritative specifications of its content." (Kisilevsky & Stone, 2019, p. 172).

[488] I am well aware that many scholars view Ripstein's reading as a form of juspositivist interpretation. Hanisch (2018), for example, maps the various interpretations of Kant's legal theory along a spectrum: at one end are scholars such as Hans Friedrich Fulda (1997), Sharon Byrd (2010), and Ernest Weinrib (2012), whose readings tend toward a quasi-Lockean view of property, contract, and status rights in the state of nature, while at the other end she places figures like Jeremy Waldron (1996), Arthur Ripstein (2009), and Alan Brudner (2011), who cast Kant's legal philosophy in a quasi-positivist light. I believe, however, that positioning Ripstein at this end of the spectrum is misleading, as it fails to account for many aspects of his reading, and especially because it does not address the fact that other scholars (Flikschuh among them) critique Ripstein's approach precisely for leaning too heavily on teleological assumptions. If Ripstein is at the right-end of the spectrum, then Flikschuh is off-chart on his right.

on the idea that institutions are justified by the exercise of pre-political natural rights. Rather, Ripstein argues that Kant constructs a conception of right and justice grounded in the principle of equal external freedom, understood as independence from the arbitrary will of others (Ripstein, 2009, p. 15).[489] Ripstein is quick to explain how this Kantian move is not based on human inclinations or vulnerabilities—features often central to natural law theory: his analysis focuses on the structure of relations among choices, rather than on empirical accounts of human nature (Ripstein, 2009, p. 7).[490]

The starting point of Ripstein analysis is, therefore, the "Universal Principle of Right", which holds that an action is right if it allows everyone's freedom to coexist under a universal law (MS 6:230). From this principle arises each person's innate right to the free specification of their ends, as long as this freedom can coexist with that of others under universal law. Kant thus draws a distinction between internal freedom, which is the *autonomy* over the law I give to myself, and external freedom, understood as *independence* from the arbitrary will of another (Ripstein, 2009, p. 14).[491] It is not merely about doing what one desires and achieving outcomes, but about not being subordinated to someone else's will in the pursuit of one's own ends and having the capacity to set one's own purposes free from others' control as a "self-master" (Ripstein 2009, 31). [492]

[489] "The person who uses your body or a part of it for a purpose you have not authorized makes you dependent on his or her choice; your person, in the form of your body, is used to accomplish somebody else's purpose, and so your in dependence is violated. This is true even if that person does not harm you, and indeed, even if he benefits you. This recasting of the familiar Kantian distinction between means and ends provides a distinctive understanding of the ways in which one person can interfere with the in dependence of another, either by drawing that person into purposes that she has not chosen or by depriving her of her means." (Ripstein, 2009, p. 15).

[490] "The direct implications that Kant draws from the Universal Principle of Right—each person's right to be his or her own master, to be presumed innocent, and to speak in his or her own name—sit well enough with considered judgments. Other, less direct implications are neither unfamiliar nor foreign. Kant understands ordinary moral thought as the exercise of practical reason, and as such, the broad structural features of familiar legal institutions will, unsurprisingly, be understandable in terms of the broad structure of practical thought. Even the less familiar aspects of Kant's arguments need to be understood in light of their relation to his austere starting point. Kant's view does conflict with what have become entrenched philosophical commonplaces, both about the anthropological nature of questions of justice and, more generally, about the appropriate concepts for practical thought. How much weight to attach to those disagreements depends at least in part on whether the familiar views turn out to be in tension with the equally well- rooted idea that no person is the master of another." (Ripstein, 2009, p. 7).

[491] "The core idea of independence is an articulation of the distinction between persons and things. A person is a being capable of setting his or her own purposes, while a thing is something that can be used in pursuit of purposes. Kant follows Aristotle in distinguishing choice from mere wish on the grounds that to choose something, a person must take himself to have means available to achieve it." (Ripstein, 2009, p. 14).

[492] "I will explain the social and interpersonal dimension of this conception of self-mastery, its relation to a system of equal freedom in accordance with universal law, and Kant's characterization of that system in terms of coercion. By making the innate right to freedom the basis for any further rights, Kant imposes an extreme demand for

Restrictions based on this principle apply equally to all, regardless of their specific goals, because they focus on the form of choice (how persons use their means and if they use it to violate your independence), not its content (Ripstein, 2009, p. 16).[493] Independence, then, involves a system of reciprocal limits where each person's freedom is defined in a way that makes equal freedom possible for all (Ripstein, 2009, p. 33).[494]

Kant's core concept of independence, understood as the condition in which no person is subject to the choice of another, forms thus the basis of right, which governs interpersonal relations.[495] It is the principle that guide the justification of legal norms and carries the

unity on his account of political justice. The rights that each person has against others must be derived from it, as must the fundamental constitutional rights that protect political freedoms and freedom of religion." (Ripstein 2009, 31).

[493] It is important to emphasize that this innate right to external freedom is fundamentally different from purely negative conceptions of freedom, which define liberty merely as the absence of external interference, regardless of its source. For Kant, the relational dimension remains central: the innate right is not a property of an isolated individual, but a normative principle governing relations among persons. This relationality concerns the structure of interactions between free and equal agents with respect to their means of action, rather than the particular ends those agents may pursue. As Ripstein brilliantly exemplifies: "What you can accomplish depends on what others are doing—someone else can frustrate your plans by getting the last quart of milk in the store. If they do so, they don't interfere with your in dependence, because they impose no limits on your ability to use your powers to set and pursue your own purposes. They just change the world in ways that make your means useless for the particular purpose you would have set. Their entitlement to change the world in those ways just is their right to independence. In the same way, your ability to enter into cooperative activities with others depends upon their willingness to cooperate with you, and their entitlement to accept or decline your invitations is simply their right to independence. Kant aims to show that independence, understood in this way, comprises a self- contained domain of reciprocal limits. The idea of a system of equal freedom both poses the problem and gives him the resources to provide a principled account of the most striking features of political life." (Ripstein, 2009, p. 16).

[494] "Kant conceives of equal freedom differently. It is not a matter of people having equal amounts of some benefit, however it is to be measured, but of the respective independence of persons from each other. Such independence cannot be defined, let alone secured, if it depends on the particular purposes that different people happen to have. One person cannot be independent of the effects of choices made by other people, except by limiting the freedom of those people. Instead, a system of equal freedom is one in which each person is free to use his or her own powers, individually or cooperatively, to set his or her own purposes, and no one is allowed to compel others to use their powers in a way designed to advance or accommodate any other person's purposes. You are independent if you are the one who decides what ends you will use your means to pursue, as opposed to having someone else decide for you." (Ripstein, 2009, p. 33).

[495] Ripstein's emphasis on the right of independence allows us to reconsider the debate between will and interest theories we discussed in chapter 1 and their relational nature in a Kantian sense (discussed also in chapter 3). The Kantian right to independence, for Ripstein, is not grounded in the pursuit of personal interests or in autonomy conceived as a life goal, but rather in the condition of being free from the arbitrary will of others. It is, fundamentally, a relational status: one is not subject to the purposes or control of another. Unlike interest-based or will-based theories of rights, which see rights as tools to protect valuable interests or individual choices, Kant sees rights as reciprocal constraints on freedom, grounded in a universal law that applies equally to all. The Kantian right to independence is not conditional on promoting any external good (like well-being or autonomy), but exists solely to prohibit domination by others—it is valid simply because it reflects a system of equal freedom for all. "This right to independence is not a special case of a more general interest in being able to set and pursue your purposes. Instead, it is a distinctive aspect of your status as a person in relation to other persons, entitled to set your own purposes, and not required to act as an

"justificatory burden", from the prohibition of personal injury to the separation of constitutional powers (Ripstein, 2009, pp. 13–14).[496] Ripstein (2009, p. 17),[497] in fact, explains how the idea of independence unfolds across three levels: Innate Right (which governs the most basic interactions in their simplest form between free persons), Private Right (which governs relations of property, contract, and status), and Public Right (which requires the establishment of a constitutional state).

Innate right is the right of humanity in one's own person, and it does not require any constitutive act to be established; individuals possess it simply by virtue of being capable of setting their own ends. Ripstein emphasizes how Kant articulates this concept of innate right through two complementary perspectives (Ripstein, 2009, pp. 17–18). First, every person has an inborn right to independence from others—no one is born as a master or servant. This innate juridical equality entitles individuals to use their own powers freely, as long as it does not infringe upon the same freedom in others. It also includes being beyond reproach, meaning individuals are accountable only for their own actions and presumed innocent unless proven otherwise. From the second perspective, instead, innate right entails the duty of "rightful honor" which for Kant means not

instrument for the pursuit of anyone else's purposes. You are sovereign as against others not because you get to decide about the things that matter to you most, but because nobody else gets to tell you what purposes to pursue; you would be their subject if they did. Thus, Kant's conception of the right to independence rests on neither of what is referred to in recent literature as the "interest theory" or "will theory" of rights. […] If rights are understood in this instrumental way, they are always at least potentially conditional on their ability to secure the underlying values that they are supposed to protect. The Kantian right to independence, by contrast, is always an entitlement within a system of reciprocal limits on freedom, and so can only be violated by the conduct of others, and its only point is to prohibit that conduct." (Ripstein 2009, 34).

[496] "This innate right leads to private right, which governs the interactions of free persons, and then to public right, which requires the creation of a constitutional state. The idea of independence carries the justificatory burden of the entire argument, from the prohibition of personal injury, through the minutiae of property and contract law, on to the details of the constitutional separation of powers. Kant argues that these norms and institutions do more than enhance the prospects for independence: they provide the only possible way in which a plurality of persons can interact on terms of equal freedom. Kant's concern is not with how people should interact, as a matter of ethics, but with how they can be forced to interact, as a matter of right." (Ripstein, 2009, pp. 13–14).

[497] "Kant develops the idea of independence in three stages. He first articulates the relation of independence in its simplest form as a constraint on interactions between persons. He calls this "the innate right of humanity" in one's own person, because it does not require any act to establish it. Instead, people are entitled to independence simply because they are persons capable of setting their own purposes. This form of in dependence is incomplete, and needs to be extended to take account of the possibility that people could have entitlements to things other than their own bodily powers. Those entitlements fall under private right, and cover the traditional categories of Roman private law, relations of property, contract, and status, which govern rights to things, to performances by other persons, and, in special cases, rights to other persons. These categories provide a complete specification of independence between interacting persons, but can only be consistently enjoyed by all in a condition of public right with legislative, executive, and judicial branches. Each of these branches in turn has further powers grounded in its role in providing a rightful condition." (Ripstein, 2009, p. 17).

allowing oneself to be used merely as a means for others' ends (Ripstein, 2009, p. 18). Though this is an internal duty (since it cannot be externally enforced), it is still part of the doctrine of right because it defines the conditions under which external freedom can be justly exercised. Consequently, individuals cannot voluntarily enter into arrangements that treat them as mere means—like selling themselves into slavery—or bind themselves in ways they could not reciprocally bind others.[498] The violation of the right to freedom (understood as independence) takes the form of domination, in which one person uses another's powers without authorization or deprives them of the means to pursue their own ends (Ripstein, 2009, p. 42).[499]

Private Right concerns relations between individuals and is grounded in the presupposition of each individual's innate right to freedom and independence (Ripstein, 2009, p. 47);[500] it introduces the concept of acquired rights, which require affirmative acts to be established (unlike innate right, which is inherent in humanity). Ripstein argues that Kant's concept of private right is structured around the categories of Property, Contract, and Status. 1) Property concerns rights over things. The ownership of an external object is the right to dispose of it for the purpose of setting and pursuing one's own ends. It includes both possession (the right to control the object and exclude others) and use (the freedom to utilize the object for one's purposes without needing others' consent). Interference with another's property—such as damage or trespass—constitutes a wrong because it either deprives the owner of the means to pursue their ends or uses their property for ends that are not their own. (Ripstein, 2009, p. 67).[501] Property is a relation among persons with

[498] Since innate right is the ground of private and public right, this principle applies also to the higher levels, restricting even the state's authority to subject individuals to another's private purposes.

[499] Ripstein aptly underlines how this perspective focuses on the act of the transgressor rather than on the harmful consequences in themselves. As he puts it: "The right to freedom as in dependence provides a model of interaction that reconciles the ability of separate persons to use their powers to pursue their own purposes. In so doing, it also provides a distinctive conception of the wrongs that interfere with this in dependence. Wrongdoing takes the form of domination. Both your right to independence and the violations of it can only be explicated by reference to the actions of others. Wrongs against your person are not out comes that are bad for you which other people happen to cause. Unlike the familiar "harm principle" put forward by Mill, which focuses exclusively on out comes that can be characterized without reference to the acts that bring them about, the right to freedom focuses exclusively on the acts of others. It is not that somebody does something that causes something bad to happen to you; it is that somebody does something to you." (Ripstein, 2009, p. 42).

[500] "The right to engage in consensual interactions and the rights you acquire through consensual interactions are, strictly speaking, not parts of the innate right of humanity as such. Instead, they are acquired rights, which require affirmative acts to establish them." (Ripstein, 2009, p. 47).

[501] "For Kant, property in an external thing—something other than your own person—is simply the right to have that thing at your disposal with which to set and pursue your own ends. Secure title in things is prerequisite to

respect to a thing, not merely a relation between a person and a thing. 2) Contract concerns rights held against specific persons. A contract allows parties to modify their reciprocal rights, so that one person acquires the right to expect a specific performance from another. Through agreement, the right to that expectation is transferred. A contractual wrong occurs when one party deprives the other of a means (their performance) to which they had given a right, thereby interfering with the other's external freedom (Ripstein, 2009, p. 69).[502] Consent is essential to contract, but it is not an independent source of moral foundation it only matters when it can render otherwise wrongful actions permissible. Exploitative actions are wrong because they are exploitative, not simply because they lack consent (Ripstein, 2009, p. 71).[503] 3) Status concerns rights over persons analogous to rights over things. It refers to situations in which a person is unable to give valid consent, as in the case of children (Ripstein, 2009, pp. 71–76).[504] A wrong related to status consists

---

the capacity to use an object to set and pursue ends. Secure title has two parts to it, possession and use. You have rightful possession of a thing provided that you are entitled to control the thing and exclude oth ers from it. Thus you are wronged if someone else damages your property, or trespasses against it. If your property is damaged, you are deprived of means you could have used to set and pursue ends. If your property is trespassed against, it is used in pursuit of ends that you have not set for yourself (Ripstein, 2009, p. 67).

[502] "If you and I make a contract, each of us agrees to do something for the other, and each of us transfers the right to expect that deed to the other. We act interdependently and consensually. Through our agreement, I do not acquire an external thing, but your deed. People may rely upon the behavior of others in a variety of ways; contract is distinctive because it creates an entitlement: you can demand that I perform, and a remedy if I fail to, because I have failed to give you something to which you have a right. Without a contract, you have not been wronged if your expectations are frustrated. In the case of a contractual right against you, I do not possess you. I possess only your power to bring about a specified outcome in the manner specified by our agreement. As a rightful relation, contract also makes a distinctive type of wrong possible. I wrong you if I deprive you of a means—my performance—to which I have given you a right. Put slightly differently, the wrong consists in my failure to advance your ends in a way that you have a right to have me advance them." (Ripstein, 2009, p. 69).

[503] "[I]t is easy to be seduced by the idea of consent, and to suppose that it is a self- standing source of moral significance. This appears to be the view of some libertarians, for example. But consent does not work that way at all. We don't worry about the lack of consent except where we are concerned with an action that would be wrongful but for the presence of consent. So if you want to know what is wrong with exploitative relationships, say, it is not that they are nonconsensual. It is that they are exploitative. It is just that consent can sometimes make a relationship in which one person determines what ends another person's means will be used to pursue nonexploitative. If an act is not wrongful, no consent is required. A wrongdoer does not need to consent to the redress to which his victim is entitled. Nor do we determine the nature of that redress by asking what the parties would have agreed to in advance. Instead, we need to ask what would right the wrong." (Ripstein, 2009, p. 71). Ripstein further outlined three key moral limits in interpersonal relations in in which one person can wrongfully interfere with another (that interference with another's person or property requires consent, that contracts must be honored otherwise they violate consent, and that one may not use another to pursue personal ends without their consent) and consent is precisely what can make that interference non-wrongful.

[504] In such relationships, one person (e.g., a parent) has possession of the other (e.g., the child) but not the right to use them for their own ends. Instead, they have a duty to act for the benefit of that person—understood as the capacity to become an autonomous being.

in using another person for one's own ends, thereby depriving them of the freedom to set their own. These three types of wrong—concerning property, contract, and status—are considered exhaustive of the private wrongs that interfere with external freedom among individuals. What allows these wrongs to be overcome is the possibility of forming a "united will" between the parties (Ripstein, 2009, p. 113).[505] By uniting their wills, individuals can legitimately redefine the boundaries of their reciprocal freedom.

Public right expands this level of interactions, since it concerns the conditions under which a plurality of individuals, each with their own private rights, can coexist under a political constitution. It is necessary because, in the *state of nature* (an hypothetical condition lacking public legal institutions), private rights—though conceptually valid—cannot be definitively secured due to three key "defects" (Ripstein, 2009, pp. 145–153).[506] Public right seeks to resolve these defects through the establishment of public institutions and a legitimate authority. The creation of a civil condition through a public constitution is the only way to transform the provisional private rights of the state of nature into peremptory rights. The united will of the people, which is the foundation of public right (Ripstein, 2009, p. 277),[507] finds its full expression through public legislation, which encompasses constitutional law, international law, and cosmopolitan law (Ripstein, 2009, p.

[505] "Kant introduces the concept of a "united will" through which the transfer from me to you is continuous because it engages both of us. Kant notes that people making agreements often try to represent this continuity by physical acts that are continuous in the same way, such as shaking hands or breaking a straw. The basic idea, however, has nothing to do with the continuity of the handshake, any more than it has anything to do with the continuity of the physical transfer of a small object. Instead, the reason the watch becomes yours is that we create an entitlement together, that the watch becomes yours. I transfer it to you, and you accept it from me. We act together, because my act of transfer and your act of acceptance each takes the other's act as its object." (Ripstein 2009, 113).

[506] The three defects of the state of nature are: 1) Judicial defect: the absence of an authority capable of issuing definitive judgments in disputes over rights. Each person acts as judge in their own case, which leads to uncertainty and potential conflict. 2) Executive defect: the absence of an authority capable of coercively enforcing rights. Even when a right is clear, there is no reliable way to prevent its violation or to obtain redress. 3) Legislative defect: the absence of an authority capable of establishing universal and binding laws. Without public legislation, rights remain provisional, and their application depends on the agreement of all parties involved. "The three defects are distinct, but have a parallel structure: nobody is under any obligation to defer to the deeds, claims, or judgments of others, unless appropriate institutions are in place. The distinctive powers that each institution must have require that those institutions differ in kind from any sort of private association. A private association can only have such powers as its particular members transfer to it. The powers to authorize one person to change the normative situation of all others, to enforce private rights in the name of all, and to impose closure on private disputes are all powers that no private person could have. The point of each argument is to show that these powers are morally required even though private persons lack them." (Ripstein, 2009, p. 147).

[507] "The condition of public lawgiving, including lawgiving that makes the acquisition of property binding on others, and lawgiving that makes rights to external objects of choice enforceable, is a united will." (Ripstein, 2009, p. 277).

229).[508] Constitutional law concerns the fundamental structure of the state, the powers of its institutions (legislative, executive, and judicial),[509] and the rights of citizens in relation to the state; international law governs relations between states and, ideally, promotes perpetual peace through a federation of free states that have transcended the need for mutual coercion; cosmopolitan law is limited to a right of hospitality and regulates relations between individuals and foreign states, rather than between individuals and the citizens of those states. In short, although public right focuses on the creation and maintenance of a legal condition, its primary aim is to guarantee the freedom and independence of each individual within a civil (national and international) society.

Now that we have outlined the basic structure of Ripstein's book, we can turn to how his argument addresses the main challenges faced by jusnaturalist reading: the relationship between moral and legal obligations and the role of the state (especially in the justification of coercion and in the possibility of a right to revolution or resistance). One of the principal merits attributed to Ripstein's interpretation is that it offers a way of reading Kant in which the foundation of law does not derive directly from the supreme principle of morality or the categorical imperative; instead, law is given an independent grounding, rooted in the notion of external freedom as independence from subjection to the arbitrary will of another (Kisilevsky & Stone, 2019, p. 145).[510] Yet, although distinct, Ripstein still stress that law is prior and sets limits on the means that agents may use,

---

[508] "The distinctive moral status of states generates each of the three "definitive" articles for perpetual peace: it requires republican governments, in which the people act as a collective body rather than a mob; a federation of free states, which have "outgrown" the need to be under coercive law; and cosmopolitan right, which is limited to a right of hospitality (as opposed to colonialism) because it governs relations between an individual and a foreign state, rather than between the individual and the individual citizens of that state." (Ripstein, 2009, p. 229).

[509] The separation of powers emerges as a necessary condition to overcome the inherent problems of the state of nature and to establish a condition of public right in which rights can be defined, applied, and enforced in a manner consistent with the freedom of all. It is essential to avoid despotism and to ensure that the state acts in accordance with the law.

[510] "Ripstein shows how Kant rejects the common assumption of many philosophers, that political philosophy (and more generally the area that Kant designates by the term Recht) is simply 'an application of general moral principles to the factual circumstances that make political institutions necessary'. In effect, with this claim Ripstein is denying that the principle of right, and duties of right, are grounded on or derived from the principle of morality. And Ripstein's book contains an admirable development of an alternative foundation for right—namely the identification of right with the conditions for protecting what Kant calls 'external freedom'—freedom as the independence from constraint by the choice of another. Considering the conditions under which people may have external freedom in accordance with universal law will permit us to see how right might be a separate sphere of morals (Sitten), entirely independent of the sphere of ethics, resting on considerations of practical reason that do not depend on the categorical imperative that grounds ethical duties, though the basis of right may have something in common with the basis of ethics, making both of them spheres of morals" (Kisilevsky & Stone, 2019, p. 145).

thereby shaping the maxims on which they can act (Ripstein, 2009, pp. 387–388). Law identifies individuals as embodied rational beings, which is the presupposition for discussions of duties toward others in the Doctrine of Virtue. In this sense, law creates the external conditions necessary for morality to be realized. Conceiving the foundation of legal obligations as distinct but not radically separated from moral obligations, is probably the main sense in which Ripstein surpass the dichotomy between jusnaturalism and juspositivist, as underlined by George Pavlakos (2017, p. 118).[511]

While some scholars like Sharon Byrd and Alan Brudner put more emphasis on the notion of innate right (Hanisch, 2018, p. 102)[512], Ripstein's reading aptly underlies how the rights in the state of nature are described by Kant as "provisional" (Ripstein, 2009, p. 165) due to the inherent defects of that condition (the indeterminacy of rights, the lack of reciprocal assurance, and the possibility of unilateral appropriation); only through public right do these rights become "conclusive" and genuinely effective. While this perspective might appear to align Ripstein more

[511] "Over the years the doctrine of right has been read both as a defence (or proof) of legal positivism and as offering a powerful case for non-positivism. Arthur Ripstein's recent reconstruction rightly seeks to place the doctrine of right beyond such crude dichotomies in suggesting that legal obligation be accounted for in terms of omnilateral willing, or the form of willing that pertains to coercive public institutions that are invested with moral authority." Pavlakos also explicitly underlined how Arthur Ripstein situate himself regarding specifically to the separation thesis. "This reconstruction has led to the claim that Kant is a positivist when it comes to legal obligation: in other words, he endorses some form of the separation thesis, or the claim that there exists no necessary connection between legal and moral obligation. What is more, under a more stringent version of the claim, legal obligations are grounded on coercive social institutions, whereas moral obligations are grounded on a pure test of reason. Ripstein rightly would like to reject the dichotomy lurking in this picture. One need not think—in Ripstein's view—that rational determination in respect of legal obligation should be conceived of in a manner that is radically disconnected from rational determination in the case of moral obligation." (Pavlakos, 2017, p. 118).

[512] However, notice that even these two scholars resist being confined within the rigid categorizations of jusnaturalism and juspositivism. Sharon Byrd thinks that the function of the state in Kant consists in a reaffirmation (with force and authority) of the previous natural rights. She writes: "Kant notes that without a right to property and other objects of our choice there would be no duty to move to the civil social order. Property rights therefore are rights we have in the state of nature. They do not depend on social approval any more than our right to freedom of choice in general depends on social approval and recognition. The sole purpose of the state for Kant is securing rights we already have before leaving the state of nature and moving to the civil state" (Byrd 2010, 94). While for some juspositivists scholars this reading has difficulties on explaining some passages where Kant explicitly says that rights are so only with a state, this is not problematic for Byrd, because Kant's legal theory concerns the rights we are entitled to claim, not merely those we possess morally. The state does not create rights, but protects them. Brudner also interprets Kant's legal philosophy as a synthesis of jusnaturalism and juspositivism, but with a distinctive Hegelian twist that leans toward institutional positivism while preserving a rational foundation (Brudner 2011, 282). According to Brudner, Kantian right is incapable of providing a freestanding, i.e., non-instrumentalist and non-public, justification for these private law institutions and their specifics. Brudner allows private right to exist but only "post-institutionally", that is, after (conceptually speaking) this legal realm has been established and authorized by the omnilateral will. All legal normativity that private law can ever unfold is due to it being embedded within a comprehensive system of public law. Hence, the strict hierarchical supremacy of the latter over the former in Brudner's interpretation

closely with a form of juspositivism, it is crucial to emphasize that the institutionalization of these rights is not an arbitrary or merely contingent process, but rather the natural outcome of the practical reasoning of rational agents.

Regarding specifically the question of the role of the state on the relationship between natural and positive rights, usually philosophers that give a jusnaturalist reading of Kant argue that the state capacity of securing rights (by creating a rightful condition and public law) does not make a normative difference (the normativity emanates only from the natural-innate rights). However, since the positions override and blur here, as in many other cases of contrasting jusnaturalism and juspositivists, it is preferable for many to distinguish only between philosophers in favor of the derivation and philosophers in favor of the non-derivation (Kisilevsky & Stone, 2019, p. 28). For example, Leslie Mulholland (1990) and Paul Guyer (2005) support the derivationist interpretation, maintaining that, for Kant, the state and its authority can be derived or justified from prior or more fundamental principles—such as individuals' innate rights or the necessity of effectively securing those rights. On this view, the state emerges as a pragmatic consequence of these principles: a practical instrument designed to ensure the conditions of freedom and well-being, both historically and contingently motivated. Others, such as Alan Wood (2014) and Martin Stone (2017), argue instead that the state is the realized form of external freedom and follows as a logical consequence of it. This second interpretation presents a more synthetic line of reasoning: the concept of the state develops internally from the very idea of freedom, evolving, like a seed, from a simple and abstract notion into a complex and articulated structure—an epigenesis of pure reason.[513]

According to Ripstein, then, law and political authority are not, for Kant, merely instruments for achieving external moral ends, but rather constitute an autonomous system aimed at securing the equal freedom of all. The legitimate use of force by the state is, therefore, intrinsic to this system itself and therefore a "legitimate type of coercion" (Ripstein, 2009, p. 55). Ripstein is very clear that if coercion is primarily understood as sanction—the threat of punishment for disobedience—we are treating coercion as a secondary element in relation to the foundational

[513] This latter reading appears more coherent with Kant's political philosophy as a whole. Notably, in *Perpetual Peace*, Kant outlines a vision of cosmopolitan right that must be extended globally (8:357–418). Were the state merely a contingent instrument for resolving conflict, Kant's cosmopolitan project would seem deeply problematic, if not altogether unfeasible.

principles of political philosophy (Ripstein, 2009, pp. 53–54).[514] Kant adopts a conception of coercion that does not reduce it to mere sanctions or threats, but rather understands it as an intrinsic limitation of individual freedom within a legal system. This conception is central to his political philosophy, which aims to define a system of equal freedom under universal laws, where coercion plays a constitutive role in ensuring respect for the reciprocal limits of freedom. Coercion is thus viewed as the removal of an obstacle to freedom (namely, the violation of another's right), rather than merely as a tool for securing obedience through the fear of punishment.

If coercion is rational insofar as it realizes the only possible system capable of securing the equal freedom of all, can one still refuse to be coerced? As we have seen, there are passages in Kant that seem to rule out this possibility. Yet the question remains: why should we not be entitled to resist a state that fails to serve the common good? According to Ripstein, two key concepts are central here: the "united will" and the idea of "contract", and they should be seen as "idea of reasons" in the Kantian sense. The concept of a united will is fundamental for overcoming the defects of the state of nature and for grounding a system of public right that can secure and conclusively determine private rights, such as property (Ripstein, 2009, p. 156).[515] Subsequently, the concept of social contract is not a historical event but an idea of reason that embodies this notion of a united will: it functions as an ideal model by which we can assess whether existing laws and public institutions can be seen as expressions of a general united will and, as such, legitimate (Ripstein, 2009, p. 325).[516]

---

[514] "If coercion is understood in terms of sanction, it must have a secondary place in political philosophy, and not figure in its basic principle, as Kant suggests. The idea that the making of threats is somehow constitutive of law or the state is vulnerable to a familiar line of objection, made prominent by H. L. A. Hart […] Instead, any adequate account of law must begin with the concept of a rule or norm, rather than trying to reduce it to the concept of a threat. Kant does not conceive of coercion in terms of threats, but instead as the limitation of freedom." (Ripstein, 2009, pp. 53–54).

[515] "Kant's appeal to the idea of a united will makes the object of agreement the rule of law through political institutions, so that individual acts of rule-making are themselves instances of a more general law. The argument is not supposed to show that an agreement has happened, or even that it would be wise or prudent for people to enter into such an agreement so that it would happen under ideal circumstances. It shows only that a form of public authorization on behalf of every one is required to underwrite private appropriation. Private property requires public right because they are both instances of a single, common problem, which has an irreducibly public element. Rather than trying to reduce the public to the private, Kant's argument shows that the private is only rightful in the context of the public"(Ripstein, 2009, p. 156).

[516] "Kant's focus on the social contract as an "idea of reason" solves a certain problem about any actual existing set of institutions. Both the entitlement and the obligation to think of the government as a representative of the united general will follow from the need to look at any sort of particular institutional realization of law as an imperfect version of the fully specified idea of law." (Ripstein, 2009, p. 325). Ripstein conception of the contract as

Understanding these two key-concepts as idea of reasons permits to see how for Kant even the most basic private rights presuppose a public authority and, therefore, "there can be no general objection to authority as such" (Ripstein, 2009, p. 326).[517] Although the state has the duty to strive for greater conformity with the idea of the original contract, the people do not have the right to use force to impose such conformity. The only recourse available to the people against oppressive legislation is through "complaints", not through resistance or a revolution (Ripstein, 2009, p. 329).[518] A revolutionary, acting unilaterally, lacks any rightful authority to employ force, since the concept of *the people* has normative significance only within a rightful legal condition (Ripstein, 2009, p. 28).[519] Outside such a condition, there is no legitimate collective will, and thus no valid title to coerce in the name of justice.

---

an idea of reason cites directly O'Neill's *Constructions of Reason* as a source of inspiration. "As Onora O'Neill has argued, Kant's focus on what the people "could" choose differs from more recent contractarian theories that focus on what people in specified circumstances "would" choose, in order to best se cure their own prospective advantage. Although the distinction between what could and what would be chosen can be collapsed by insisting that people could only choose what is most advantageous for them, Kant's emphasis on possible choice focuses on the grounds for the individuals accepting the authority of the state, that is, to guarantee the systematic enjoyment of the right to freedom. You could agree to restrict specific exercises of your freedom, in order to guarantee its systematic preconditions, but the prospect of advantage could not entitle you to "throw away your freedom" in ways inconsistent with your general right to be your own master. No such act could be an instance of protecting your own capacity for choice." (Ripstein, 2009, p. 206). We will return on the relation with Ripstein and O'Neill at the end of this sub-section, but we should also briefly mention here how they disagree in an important sense. While we said previously that Ripstein sees the universal principle of justice as distinct but related to the categorical imperative, O'Neill will later express a more direct subordination. "Put schematically, Kant's justification of the social contract as an idea of reason relies on three linked claims. Proceeding from the most to the least abstract, these claims are: A. The categorical imperative is the supreme principle of practical reason, so an idea of reason. B. The universal principle of justice states a version of the categorical imperative, restricted to the public domain. C. The republican conception of the social contract is a special case of the universal principle of justice, adjusted to certain historical conditions" (O'Neill, 2015b, p. 180).

[517] "Kant's insistence that citizens must obey their governments and have no right to take up arms against it is often said to be objectionable, and is sometimes said to be shown by experience to be objectionable. We have already seen Kant's systematic answer to the philosophical anarchist, who doubts that authority could ever be legitimate, and so doubts that an official could ever be entitled to make a choice that binds others. Even the most basic private rights presuppose a public authority. So there can be no general objection to authority as such, or to institutions vesting authority in individuals." (Ripstein, 2009, p. 326).

[518] "Kant's arguments against revolution rest on the difference between the postulate of public right and the idea of the original contract: a state is under a duty to bring itself more nearly into conformity with the idea of the original contract, and the people are "authorized at least to make representations against" laws that are not in conformity with it. The state's internal duty does not correspond to a correlative right on the part of the people, which is just to say that the people may not use force to bring the state into conformity with it. If the state violates this duty, the people do not have an enforceable remedy, and so must "put up with" oppressive legislation. Their only recourse is to "oppose this injustice by complaints but not by resistance." (Ripstein, 2009, p. 329).

[519] "[T]he revolutionary does not claim to be speaking for himself, but rather to be acting on behalf of "the people." Yet the idea that there is such a thing as a people only has application in a rightful condition. A rightful condition solves the problem of unilateral choice and action by creating institutions through which the people can act

Nonetheless, while a right to revolution is explicitly denied in order to avoid a logical contradiction within the legal framework, this does not entail subjugation to any and every form of government. The state is necessary to overcome the defects of the state of nature (the lack of mutual assurance of rights, the indeterminacy of law, and the unilateral character of property acquisition) and to establish a rightful condition in which each person's freedom can coexist with that of others under universal laws. However, if the state itself becomes a source of insecurity for rights, acts arbitrarily, or imposes obligations incompatible with equal freedom, it fails in its justificatory purpose. In such circumstances, the very reason why opposition was considered problematic—namely, the risk of reverting to the defects of the state of nature—may be weakened, since the state itself would be creating or perpetuating those very defects. If the state enacts laws that the people could not give to themselves while preserving their freedom and innate rights, it acts in contradiction with this idea of reason (Ripstein, 2009, p. 209).[520] While it is important to recognize that, in Kant's philosophy, a right to revolution is explicitly denied due to a logical contradiction within the juridical framework, another central issue pertains how we might recognize and respond when a state fundamentally departs from the principles that justify its authority.

Overall, as we have seen, Ripstein's interpretation has been described as either a form of jusnaturalism or juspositivism, depending on the critics. However, the most substantial critique is arguably directed at his jusnaturalist leanings—particularly his strong reliance on teleological

---

as a collective body. The most that the revolutionary can claim to do is speak unilaterally for some end or other. His unilateral claim, however, could never authorize the use of force. Thus, the revolutionary has no right, that is, no title to coerce. Understood in this way, Kant's arguments against revolution do not, however, lead to the conclusion that some of the horrible regimes that the world has seen must not be resisted. To the contrary, the most egregious of these are not even candidates for a united will, because their founding principles, such as slavery or genocide, do not create a rightful condition in which all can be secure in their rights." (Ripstein, 2009, p. 28).

[520] For example, a law instituting slavery or unjustifiably suppressing freedom of expression—beyond what is necessary to protect the rights of others—would violate the foundational principles that justify the existence of the state. "Kant is concerned with the authorization to coerce, so it cannot matter whether someone would foreseeably lack an internal incentive to conformity with positive law. If something is wrongful, it can be prohibited, no matter how significant the contrary inclinations. As Kant observes in his discussion of the so- called right of necessity, the fact that in sufficiently dire circumstances wrongful conduct could not be prevented by externally given law does not make that conduct rightful or place it be yond the scope of legal regulation, even if it makes punishment pointless. The foreseen ease or difficulty of compliance depends on the matter of choice. The argument against institutional slavery cannot be that rational persons would not take on a burden they expect to have difficulty in meeting, because questions of right are never questions about burdens at all. Instead, a rightful constitution could not institute slavery because it cannot make arrangements between private persons that those persons could not be entitled to make for themselves. If a person cannot bind him or herself to a condition of slavery, neither can an official bind that per son. No expected material advantage can override this." (Ripstein 2009, 209).

concepts and the prominent role he assigns to private law. Katrin Flikschuh partially concurs with these concerns and, in response, reorients her reading of Kant to place greater emphasis on the concept of public right. For Flikschuh, Ripstein's emphasis on innate right as an affirmation of each person's innate domain of external freedom—including the powers of control over one's own body—and his subsequent analysis of property rights (acquired right), could lead to interpreting acquired right as merely an extension of innate powers of control over the external world (Flikschuh, 2017, pp. 63–64).[521] Flikschuh argues that Ripstein's characterization of public politics as aimed at securing a "regime of equal private freedom" mirrors his interpretation of the innate right to freedom as a "system of equal freedom" (Flikschuh, 2017, p. 64); this approach, she fear, could lead to an overly disaggregated view of public politics (Flikschuh, 2017, p. 65).[522]

The relationship between natural and positive, private and public and provisional and conclusive, should be expressed differently. Against Ripstein's emphasis on "purposiveness" (the capacity to set ends) as a core element of innate right, which express " a conception of innate Right as first or founding principle", Flikschuh (2017, 60) offers an alternative reading of innate right as merely a claim to juridical equality with respect to all others, without attaching to it any substantive rights, departing from Ripstein's emphasis on powers and innate capacities.[523] This

---

[521] "By the end of chapter 2 of Force and Freedom, a dominant impression has emerged of the innate right to freedom as affirming each person's innate domain of external freedom, including, in particular, each person's powers of control over her own body and associated physical and mental capacities and talents. With respect to this innate freedom domain each is sovereign, such that others' non-consensual interference with another's body and powers constitutes a violation of her innate right. The ensuing analysis of property rights—acquired Right—now looks set to tackle the question of how persons might rightfully extend their innate albeit highly limited domains of external freedom to cover rights in things that are acquired, ie not possessed innately. I shall not here consider Ripstein's account of property rights; I want merely to note that Ripstein's ambivalent approach to innate Right impacts on his analysis of acquired Right in a manner he may not intend: given the focus on unhindered use of innate powers in the discussion of innate Right, readers may easily be moved to interpret acquired Right as seeking to extend innate powers of control over the external world and not, as I think Ripstein intends, as illustrating the morally problematic effects of our choices on others." (Flikschuh, 2017, pp. 63–64).

[522] "According to Ripstein, therefore, the Kantian state envisages a mode of public policy-making that is focused fundamentally on the creation and re-creation of a regime of equal private freedom. This is a highly disaggregative approach to public policy—one which may remind recent generations of British citizens of the Thatcherite slogan that there is 'no such thing as society'." (Flikschuh, 2017, p. 65).

[523] Flikschuh, however, is also quick to recognize that there are not foundationalist drifts in his reading. As she puts it: "Later on in the same chapter, Ripstein explicates the innate right of each with reference to persons' power to set and pursue whatever ends they see fit, seemingly grounding it in the rights-creating individual capacity for 'purposiveness'. However, Ripstein is not unambivalently foundationalist about it. While the innate right to freedom, understood as each person's independence from another's arbitrary power of choice, is said to 'carry the justificatory burden of the entire argument', Ripstein interprets the independence requirement itself relationally, not attributively.

seems to drift toward a liberalism of private freedom, which Flikschuh sees as a misreading—or at least a superficial interpretation—of Kant's thought. According to her, Ripstein may fail to grasp the full depth of Kantian legal theory, which views the juridical distinction between "mine and thine" as politically mediated.[524] No right is entirely natural: innate right describes only the a priori form of determinable rights, not "substantive entitlements" that preexist the public legal order.

Yet, despite having significantly different interpretations of the relationship between innate right and public right,[525] both authors nevertheless endorse a form of non-derivationism regarding the state, in which claims to external freedom serve as a means to establish the state, rather than viewing the state merely as an instrument for protecting preexisting rights. Similarly, Flikschuh clarifies, like Ripstein, that Kant does not support a right to revolution within a lawful state and she shows that the denial of such a right in a juridical condition does not imply passivity in the face of regimes of barbarism that exist outside any legitimate legal structure (Flikschuh, 2008, p. 396).[526]A legitimate legal condition requires a unified "omni-lateral" general will as a necessary condition for the assertion of legal claims that are valid yet unenforceable by individuals

---

Kantian independence is not a good to be promoted, but a moral claim of a kind that 'constrains the conduct of others'. Persons have an innate right to freedom against one another, not absolutely" (Flikschuh 2017, 60).

[524] The fact that "mine and thine" are politically meditated achievement is also stressed extensively by Japa Pallikkathayil's critique of Ripstein (Kisilevsky & Stone, 2019, pp. 16–17).

[525] Ripstein responds to Flikschuh by arguing that her conception of innate right as purely formal and not empirically instantiable eliminates the distinction between theft and assault. He challenges Flikschuh's view that the non-consensual moving of someone's body is a matter of "public policy making" rather than a violation of individual rights, pointing to a significant divergence in their understanding of innate right and its relationship to public authority. "This appears to leave no space for the idea of wrongs against the person. By conceiving innate right as merely prepara tory for public right, Flikschuh not only has difficulty explaining Kant's con trast between dragging someone's body from its resting place and taking that person's land; she also appears to reach the conclusion that the wrongfulness of the former is a question of 'public policy making', a matter for a public authority to decide, rather than the rights of the parties. Thus she supposes that in dealing with it, a public authority may not be 'preoccupied with ques tions about individual freedom'. However one might finally evaluate such a view, it is not Kant's position. Your innate right is a right to what you have already, prior to any act." (Kisilevsky & Stone, 2019, p. 192).

[526] "Where oppression has reached a level of intensity so as to become literally unbearable, social violence may finally erupt and dissolution result. However, far from constituting a right of the united people against its head, violent dissolution signals the complete breakdown of subjects' and sovereign's respective attempts at establishing relations of Right. Such failures of state happen- nature (in current parlance, 'circumstance') lends a helping hand, enabling those involved to attempt the civil union anew. Kant does not deny the possibility of revolution as social event; nor does he deny the possible progressive political consequences of such events. He denies only that there can be a right to revolution. The elementary reason for this denial lies in his conception of the public character of Right and in the consequent constitutive unfitness of unilateral as private wills to take public law into their own hands." (Flikschuh 2008, 396).

(Flikschuh, 2008, p. 393).[527] For both authors, therefore, the establishment of a civil condition is the necessary outcome of practical reason, without this necessarily implying an overly teleological reading rooted in natural law theory. Law is reason's public embodiment—neither divine command nor sovereign whim, but freedom's institutional mirror.

Against the interpretations that frame Katrin Flikschuh as offering a juspositivist reading of Kant, and Arthur Ripstein as providing a jusnaturalist interpretation, I hope to have clarified in what sense both authors converge in recognizing that Kant offers an alternative to both traditions, by reconciling them in the reconciliation of ethics and law. Kant, so interpreted, presents a perspective in which innate right (or its formal structure in Flikschuh's sense) both requires and is realized through public right, whose purpose is to establish and safeguard relations of equal external freedom. This process is grounded in principles that are neither merely given by nature in a substantive sense (as in traditional natural law theory), nor solely based on social convention or authoritative enactment (as in traditional positivism); rather, public right plays a necessary role in rendering rights determinate and coercible on the basis of a priori principles of legitimate external freedom.

This is also the view that any Kantian constructivists scholars will endorse. It is, for example, the view defended by Onora O'Neill, who agrees with Ripstein in emphasizing the Universal Principle of Justice at the heart of Kant's thought, and with Flikschuh in critiquing contractarian or foundationalist accounts of political authority. In the book *Constructing Authorities: Reason, Politics and Interpretation in Kant's Philosophy* (2015), and especially in the paper "Kant and the social contract theory" included in it,[528] O'Neill offers a constructivist interpretation of Kant in order to stress the need for a public legal framework grounded in rational imperatives. Her analysis centers on the role of the state and the concept of the social contract,

[527] "In Kant's pre-civil social condition, persons unavoidably raise valid entitlement claims against one another. Given their reciprocally coercive character, these claims cannot permissibly be enforced against one another in that condition. Only an omnilateral will has authority to enforce reciprocally coercive entitlement claims against all. This omnilateral will is the idea of the 'collective general (common) and powerful will' (VI 256). Despite its Rousseauean inspiration, Kant's 'collective will' does not represent the idea of popular sovereignty - the thought is not that each subjects himself only to laws which he has co-authored with all others. The general united will represents the idea of a public juridical authority as necessary condition of the possible vindication of the valid but privately unenforceable entitlement claims of each." (Flikschuh 2008, 393).

[528] "*Kant and the Social Contract Tradition*" was first published in *Kant actuel: Hommage à Pierre Laberge* (Duchesneau et al., 2000); a slightly revised version appears in "*Kant's Political Theory: Interpretations and Applications"* (Ellis, 2015), where there is also a paper from Ripstein.

which she argues should not be understood as a historical event or a personal obligation, but, like Ripstein, as an "Idea of Reason*"* (O'Neill, 2015b, p. 176)*,* a rational ideal based on the universal principle of justice, which serves as a normative standard by which all existing political arrangements can be evaluated, according to how closely they approximate its demands (O'Neill, 2015b, p. 171).[529]

O'Neill emphasizes that Kant diverges sharply from traditional contractarian theories of consent by denying, respectively, 1) explicit consent (too unreliable to serve as the foundation for universally binding laws), 2) tacit consent (too indeterminate to carry justificatory force), and 3) hypothetical consent (it either depends on idealized conditions that real agents cannot satisfy, or it presupposes the very legitimacy it seeks to justify).[530] Abandoning these three consent theories, Kant introduces what O'Neill calls a modal criterion—a standard focused on the conditions of possibility for universal consent (O'Neill, 2015b, p. 178).[531] In other words, legitimacy arises from what all could rationally will, not what they empirically do or might desire. This approach reflects the broader structure of Kant's critical philosophy, where freedom becomes the central concern, and where the very structure of judgment—comparable to that of a court—is inherently political. Political authority is thus grounded in reason, not in (social) agreement.

Unlike modern contract theories that focus on what individuals would choose for their own benefit, Kant, as O'Neill shows, is concerned with what a people *could rationally will* in order to guarantee the freedom of all. For this reason, political legitimacy for Kant is not grounded in

---

[529] O'Neill lucidly stress many times in this text how the universal principle of justice is the first and unconditional duty in any external relation between individuals who inevitably affect one another. He writes for example that "Kant's basic justification of political institutions appeals not to the Social Contract but to the 'Universal Principle of Justice (Right)'. This principle is stated at the beginning of section 2 of Theory and Practice and in the opening passages of The Doctrine of Right: 'Any action is right if it can coexist with everyone's freedom in accordance with a universal law, or if on its maxim the freedom of choice of each can coexist with everyone's freedom in accordance with a universal law' (MM 6:230; cf. TP 8:289)." (O'Neill, 2015b, p. 171). She also further stresses this point by stating that: "Kant states his views on the social contract most clearly in part 2 of Theory and Practice. The section begins with the assertion that the social contract is completely distinctive and is the "unconditional and first duty in any external relation of people in general, who cannot help mutually affecting one another" (TP, 8:289)." (O'Neill, 2015b, p. 176).

[530] Kant is usually mistakenly associated especially with this third type of consent (O'Neill, 2015b, p. 177).

[531] "[Kant] suggests that it is possible agreement that is decisive. He appeals to possible universal consent, not to hypothetical consent. The criterion is apparently modal and not hypothetical." (O'Neill, 2015b, p. 178).

consent, but in the principles of practical reason and the universal principle of justice, which together constitute the first and unconditional duty in any external relation between persons who inevitably interact. If we want to speak at all of social contract in a Kantian sense, we have to understand it not as pact, but as a regulative ideal that guides lawgivers in shaping institutions that respect and realize universal freedom. Similarly, if we want to speak at all about jusnaturalism and juspositivism inside the Kantian framework, we have to understand them through their common foundation in practical reason.

4.1.4 The contemporary conception of Juspositivism

As should now be clear, classifying all the philosophers considered thus far—from Aquinas to Kant—strictly within the framework of jusnaturalism as is nowadays conceived, is highly problematic. Yet, this is precisely what some authors, the juspositivists, have done, in order to distance themselves from specific issues associated with that approach. As emerged in the previous discussion, the relationship between juspositivism and natural jusnaturalism is largely defined by opposition, with juspositivists arguing that jusnaturalism, in its most general form, upholds two distinct but interrelated theses: (1) both natural law and positive law exist, and (2) natural law is axiologically superior to positive law.[532] Juspositivists, by contrast, reject the first thesis outright (and consequently the second), asserting that only positive law exists. In this final sub-section, we will complete our historical reconstruction by analyzing more closely some of the classical jusnaturalists, evaluating the coherence and limits of this tradition.

The emergence of juspositivism as a distinct and explicit current is traditionally placed in the early 19th century, with Jeremy Bentham recognized as its central figure.[533] Bentham indeed

[532] Nonetheless even juspositivists are open to admit that some versions of natural law theory do not fully endorse the second thesis or attempt to explain, in different ways, the complex relationship between natural and positive law. Although this is especially true for some recent thinkers, such as John Finnis and Ronald Dworkin, it was also true for some contemporaries of the early positivists, as they themselves acknowledged. Lon Fuller is in this regard a notable example, since he emphasized the procedural dimension of law, arguing that the legitimacy of a legal system depends on adherence to principles such as coherence, consistency, and transparency. In his view, the moral quality of law derives from the manner in which it is enacted and applied, rather than from a hierarchical structure where natural law always supersedes positive law. Positive law, according to Fuller, can be valid as long as it respects procedural morality, even if its content does not fully align with natural law (Fuller, 1978).

famously dismissed natural law theory as "nonsense upon stilts", a phrase that encapsulates his strong rejection of the idea that inherent, universal principles of morality and justice exist beyond human-made laws (Bentham, 1789). He criticized natural law for relying on vague, speculative reasoning, lacking empirical evidence or practical utility and advocated for a legal system based on observable human behavior and social utility, grounding his normative theory in utilitarianism: the primary criterion for evaluating legal norms should be their capacity to maximize happiness for the greatest number of people.

Hans Kelsen further developed legal positivism with his Pure Theory of Law. Kelsen emphasized the autonomy of law from morality and politics by concentrating on the internal coherence and consistency of legal norms (Kelsen, 1967). He proposed a hierarchical system of norms where lower norms derive validity from higher norms, culminating in a basic norm (Grundnorm) that establishes the validity of the entire legal system. Kelsen argued that the relation between lower and higher legal norms is not one of deduction or inference. Rather, the creation of new norms under the auspices of, or using the powers granted by, the 'higher' norms of the Constitution is a matter of discovery rather than logical deduction.

Kelsen's theory seemingly led to a circular definition: a norm is a Basic Norm if all other norms in the system are created under its authority. This circularity arises because recognizing the Basic Norm as such presupposes accepting its authority. H.L.A. Hart proposed a solution to Kelsen's circularity by introducing the concept of rule of recognition, that we already discussed in the paragraph 1.1.2 of chapter 1. At the foundation of every legal system lies an actual customary norm practiced that arises from the behaviors and beliefs of legal officials. This makes Hart's theory more attuned to the actual workings of law within a society by acknowledging that legal systems are sustained not by abstract norms and its "pure normativity", as in Kelsen, but also by the real-world acceptance and recognition by those who enforce and interpret the law. A rule is legally valid and therefore law, if it conforms to the requirements set forth in the rule of

[533] Juspositivism has of course a longer history and a broader range: it has antecedents in ancient political philosophy and is discussed (and the term itself introduced) in mediaeval legal and political thought (Finnis, 1999). The modern doctrine, however, owes little to its predecessors, as most juspositivists prefer to reference only to Hobbes and Hume as influences.

recognition, which in itself can neither be valid nor invalid: the proof of its existence is in the social fact of its acceptance. [534]

Moreover, the rule of recognition is dynamic, as opposed to Kelsen's hypotetical grundnorm which is static. As society and its legal practices evolve, the rule of recognition can adapt to new norms, principles, or interpretations. This flexibility makes Hart's theory more robust in dealing with changes in legal systems and even with legal pluralism: different legal systems, or even subsystems within a broader legal framework, can have their own rules of recognition. For example, a society might recognize both state law and customary law as valid, and Hart's theory can easily accommodate such a scenario, while Kelsen's model struggles with this multiplicity. Nevertheless, the rule of recognition has faced challenges from many scholars, most notably Dworkin, as discussed earlier in paragraph 1.1.3,[535] and even within the ranks of juspositivism some scholars find Hart's resolution of the authority of law unsatisfactory. Among these is Joseph Raz, who, while recognizing the importance of Hart's rule of recognition, offers a more refined understanding of how legal systems operate.[536]

Raz does not reject the rule of recognition but downplays its role as the sole foundational element of a legal system: while essential, it provides only necessary but not sufficient conditions for inclusion within a legal system (Raz, 1979, pp. 95–97).[537] Instead, he emphasize that legal

[534] Before Hart, Alf Ross had already critiqued the Grundnorm, not for its circularity but for its redundancy, arguing that it is unnecessary for understanding a legal system, the validity of which derives from its actual effectiveness within society. According to Ross legal system is considered valid and recognized if its norms are actively applied and followed in practice. This realist approach emphasizes the practical functioning of law, rather than depending on an abstract, normative foundation like the Grundnorm. For Ross, legal validity is entirely explained by a system's effectiveness, making the presupposition of a foundational norm superfluous (A. Ross, 2019).

[535] By highlighting that judges often disagree on the criteria for determining what constitutes a valid law, Dworkin suggests that the nature of law involves more than just a rule of recognition; it also requires the moral interpretation of certain principles. This insight has been taken more seriously by legal positivists like Jules Coleman, as will be discussed on the following page.

[536] We will return to Raz's position in detail in section 4.2.2, where we will examine the specific problems associated with his account. For now, our aim is to provide a brief historical overview to illustrate how Raz's work builds on earlier juspositivist scholarship and how it engages with the positions of other contemporary juspositivists.

[537] "There is no reason to suppose that the rule of recognition refers to all the criteria of validity of a legal system, and it is clearly wrong to think that it determines them all. A criterion of validity is a set of conditions set by law, satisfaction of which is sufficient for being a law of the system. All the laws of a legal system, except the rule of recognition the existence of which is a matter of social practice, are valid; they exist in the system because they satisfy some criterion of validity. Besides the rule of recognition, other laws can also set criteria of validity. All the laws conferring legislative powers, for example, determine criteria of validity; so also does a law stipulating conditions that a social custom must fulfil to be legally binding.88 That the rule of recognition sets up some criteria of validity is

systems are better understood through the authority and functions of legal institutions and the norms they create. Raz's conception of authority, known as the "Service Conception," argues that law “serves” individuals by guiding their behavior, especially in situations where they may lack the necessary knowledge, expertise, or coordination to make informed decisions independently (Raz, 2009, pp. 136–137). Yet, for law to fulfill its purpose, individuals must regard legal norms as exclusionary reasons for action. In this view, people are expected to obey legal norms simply because they are legal norms, rather than because they align with moral principles. Legal authority is fundamentally about guiding behavior rather than providing moral truths.

Raz is concerned with preserving the autonomy of law as a distinct social institution: conflating law with morality can lead to confusion and undermine law's role as a practical guide. While he acknowledges that law can serve moral purposes, he contends that this is a contingent characteristic of specific legal systems rather than a conceptual truth about the nature of law. A legal system may aim for justice or morality, but this does not imply that morality serves as a criterion for what constitutes law. This is Raz’s “Sources Thesis”, which holds that the existence and content of law can be determined solely on the basis of social facts, without requiring reference to moral reasoning (Raz, 1979, p. 53, 1985, p. 296).[538] Thus, a law can be immoral without losing its status as law. As Raz puts it: “Two innocent truisms about the law lie behind much of the difficulty we have in understanding the relations between law and morality. The law can be valuable, but it can also be the source of much evil. Not everyone agrees to these truisms, and there is nothing inappropriate in challenging them, or examining their credentials. They are, however,

clear. There must be in every system some criteria of validity that, although legally binding, are not legally valid, hence they must be set in the rule of recognition. There is, however, no reason to think that all the criteria must be stipulated in that rule. The fact that all the criteria of validity are determined in laws that, directly or indirectly, are valid according to criteria determined in the rule of recognition, guarantees that by imposing a duty to apply the laws satisfying its criteria of validity, the rule of recognition imposes a duty to apply all the laws of the system. Furthermore, there is no reason to suppose that every legal system has just one rule of recognition. It may have more.” (Raz, 1979, p. 95).

[538] Raz defend the Sources Thesis in light of the other two possible thesis on the relationship between law and morality “Three theses with clear implications concerning the relation between law and morality have been defended in recent years. They can be briefly, if somewhat roughly, stated as follows: The Sources Thesis: All law is source-based. The Incorporation Thesis: All law is either source-based or entailed by source-based law. The Coherence Thesis: The law consists of source-based law together with the morally soundest justification of source-based law” (Raz, 1985, p. 295). Dworkin is identified with the coherence thesis, while versions of inclusive juspositivism are associated with the incorporation thesis.

truisms in being taken by most people to be obviously true and beyond question." (Raz, 2003a, p. 1).

While acknowledging the interactions between morality and law,[539] Raz's philosophy is thus in the end a form of exclusive juspositivism: legal systems must be capable of excluding moral reasoning from the criteria of legal validity, as this exclusivity is essential for ensuring that law functions as an authoritative and practical guide. In contrast, Jules Coleman advocates for a form of juspositivism that permits the incorporation of moral considerations into the legal framework, often referred to as inclusive juspositivism (Coleman, 2003).[540] Coleman builds upon and extends Hart's ideas regarding the rule of recognition, positing that moral principles can be recognized as legal norms under certain conditions—particularly in constitutional democracies where courts may appropriately employ moral reasoning in their judicial decisions.

The ongoing debate between inclusive and exclusive positivists, which remains prominent in contemporary legal theory, demonstrates that the tension between morality and law is far from resolved. Some juspositivists continue to advocate for a strict separation, drawing primarily from the early work of Hart, while others argue for a more integrated approach that acknowledges the role of moral principles in legal systems. [541] Moreover, inside contemporary juspositivism, there's

[539] "While there are necessary connections between morality and how the law is, the more significant necessary connections relate to the evaluative perspective which informs our thinking of how the law ought to be, rather than how it is. It may be thought that the thesis that by its nature the law ought to be moral is empty or trivial, for everyone and everything ought to be moral. But that is not so. To be sure, nothing should be immoral. But it is not the case that the University of Oxford, or the city of Oxford, ought to be moral in the way that the law is. The intrinsic virtue of a university, i.e. what makes a university into a good university, is excellence in learning, research and teaching. The intrinsic excellence of a city may be comfort, and the provision of certain services. What makes the law different, what makes its intrinsic excellence a moral excellence, is that it is a structure of authority of telling people what they must do. Necessarily, the law claims to have legitimate moral authority over its subjects. Hence its intrinsic virtue is to have such authority." (Raz, 2003a, p. 16).

[540] Before Coleman, other scholars who explicitly addressed the issue and introduced the term into the debate were Waluchow (1994) and Kramer (1999). In a more recent paper, Coleman (2009) also distances himself from specific Hartian notions that had characterized his earlier formulation of Inclusive Legal Positivism, such as the centrality of the rule of recognition and the endorsement of the separability thesis.

[541] The distinction between a "first" and "second" Hart, particularly after his engagement with Ronald Dworkin, has become widely recognized among scholars. While Hart did not fundamentally abandon the core tenets of his juspositivism, his later writings reflect a greater openness to a more nuanced understanding of the relationship between law and morality, the role of judicial discretion and principles, and the nature of legal obligations. Scholars who wish to uphold a strict separation between law and morality often focus exclusively on the "first Hart," downplaying or disregarding the developments in his later work.

also an ongoing tension between those who advocate for a more analytical approach to law, and those, like Brian Leiter, who argue for a naturalistic or empirical approach to understanding legal system, particularly judicial decision-making (Leiter, 2007).[542] The lack of consensus on how to address these issues, and others,[543] within the positivist framework has revitalized modern versions of natural law, such as those proposed by John Finnis (2011) and Robert P. George (2004), who vigorously reaffirm the classical principle that "lex injusta non est lex et virtutem obligandi non habet" (Finnis, 2011, p. 360). However, these approaches reintroduce some of the foundational problems previously discussed, leading the debate into what appears to be a vicious theoretical impasse. With this brief historical reconstruction in mind, we can now delve thematically into these specific issues facing both currents, starting again with jusnaturalism.

4.2 ENDURING PROBLEMS

4.2.1 Limitations of Jusnaturalism: Foundationalism and Arbitrariness

If we identify the core of the jusnaturalist doctrine, as many scholars do, in the key distinction between natural and positive law—with the former taking hierarchical precedence—some unavoidable questions arise. Two of them specifically appear of crucial importance, namely how can we demonstrate that such a natural law exists in the first place, and how can we claim that everyone universally recognizes it, given the widespread moral and political disagreement. The first question relates to the problem of queerness—how moral properties should be considered distinct from physical properties, and why. The second question connects to the problem of

[542] Leiter's naturalized legal positivism might seem to evoke a blending of the two categories of our historical reconstruction, but it does not refer to jusnaturalism. Rather, it represents a fusion between juspositivism and scientific naturalism—the idea that theoretical philosophical reflection and experimental scientific inquiry should proceed in parallel. Nevertheless, this raises additional concerns about the ambiguity surrounding the term "natural" in legal and human rights discourse.

[543] One interesting issue worth highlighting, especially given our focus on the rule of recognition, is how legal positivists should respond to the growing complexity of legal systems due to the rise of international law, transnational legal frameworks, and multilevel governance. Applying traditional positivist principles to globalized legal orders—where clear rules of recognition may not always be present—presents a significant challenge that contemporary legal positivists must address.

relativity, the well-documented anthropological reality of varying moral and political codes across cultures.

Both of these challenges were famously brought to prominence by Mackie in his critique of moral realism (Mackie, 1977). The problem of queerness is divided into two parts: metaphysical and epistemological. The metaphysical problem argues that, if objective moral facts were to exist, they would be entities of a very peculiar kind, entirely different from anything else in the universe.[544] But given the metaphysical oddity of such entities, how could we possibly know anything about them? It would require some form of moral perception or intuition—something radically different from the way we know anything else (epistemological problem). Mackie's solution is to explain ethical behavior without invoking the existence of moral facts—those strange, intrinsically motivating entities. This is further argued by acknowledging that the ontological queerness of moral facts would also necessitate accepting an epistemological strangeness, requiring a special faculty beyond our natural perception to access these moral truths. Thus, in accordance with principles of metaphysical parsimony and Ockham's razor, Mackie concludes that we should simply discard the notion of moral facts as superfluous (Mackie, 1977, pp. 38–41).[545]

---

[544] Constructivists (Darwall, 1990; Korsgaard, 1996b) but also many other philosophers (Brink, 1992; Foot, 1972), have challenged Mackie's argument against ethical objectivism by questioning the hidden assumption, derived from Hume, that reason is entirely passive and incapable of serving as the source of active moral principles, such as conscience or a sense of morality (Hume, *Treatise*, 3.1.1.11). This assumption underpins the claim that cognitive judgments can be conceived as beliefs aimed at truth, representing properties or facts, and hence, guided by the world. Constructivists approaches conceptualize moral reasoning differently as we saw (2.2.3) so that "speculative reason is not merely "recognitive", that is, meant to track things down, as suggested by the traditional metaphors of "vision" and "insight", but generative and productive" (Bagnoli, 2022c, p. 439). For further explorations on moral knowledge and moral facts according to constructivists, see also the next note.

[545] Constructivism similarly takes the demand for metaphysical parsimony seriously but avoids eliminating moral facts altogether. Instead of discarding a crucial element of our practical experience, constructivists account for moral facts by grounding them in the framework of practical reasoning, situating them within the "fact of reason" itself (Bagnoli, 2013a). It is helpful, however, to distinguish Mackie's position from strict eliminativism (Garner, 1990) and fictionalism (Joyce, 2001) interpretations, which have developed based on his error theory. Mackie aligns with moral realists in affirming that moral language is cognitivist—that is, moral statements like "murder is wrong" aim to describe objective moral facts. The issue, for Mackie, lies not in the act of moral judgment itself but in the fact that these judgments are systematically false because the objective, prescriptive moral facts they aim to describe do not exist. As he succinctly stated: "The assertion that there are objective values or intrinsically prescriptive entities or features of some kind, which ordinary moral judgment presupposes, is, I hold, not meaningless, but false." (Mackie, 1977, p. 40). Constructivism, by contrast, preserves the concept of moral facts by *de-reifying* them: moral facts are not entities in the world that moral judgments refer to, but are principles constructed through practical deliberation. They are thus grounded as "facts of reason", not as mind-independent features of the external world (Bagnoli, 2013a).

This queerness epistemic problem seems to lead to an ineliminable problematic feature of jusnaturalism, that of foundationalism, an epistemic orientation which holds that certain beliefs serve as the ultimate foundation for all other justified beliefs. [546] Foundationalism as a long story that date back to at least the ancient Greeks, and it is mostly based on the regress problem. [547] The regress problem arises from the idea that a claim (p) is justified by an argument from other premises (q and r); but these premises also need justification, which either leads to an endless chain of arguments, a loop back to the original claim, or terminates at unsupported premises. The problem thus is whether justifications rest on unprovable foundations, go on forever or circle back.

These are, in fact, the three main positions on the problem: 1) Foundationalism, the position that some beliefs do not need further justification and serve as the foundation, ultimate premises, for all other beliefs. Foundationalists argue that this is the only way to avoid both infinite regress and circular reasoning (Alston, 1976).[548] 2) Coherentism, the position that justifications form a

[546] We have mentioned foundationalism as one of the main pitfalls that constructivism seeks to avoid. However, it is important to clarify that what constructivism primarily rejects is ontological foundationalism—that is, the idea that moral principles are grounded in some independent ontological reality. This is distinct from epistemological foundationalism we are focusing on now, which is concerned with the epistemic justification of beliefs or rational guarantees. In this sense, epistemic foundationalism is a position that does not necessarily preclude constructivism, and constructivists positions like that of Forst might as well be considered also (epistemologically) foundationalist. Yet, calling all theories that solve the regress problem "foundationalist" risks hiding the distinction between external (ontological) and internal (constitutivist) responses to the regress. This is why many constructivists prefer to preserve the distinction between foundationalism and constructivism as the main contraposition, because it allows us to say that there are two clearly distinct ways of resolving the (epistemic) issue of regress: one through an ontological external move (pointing to moral facts or properties), and the other through a move internal to reasoning itself (by identifying constitutive norms of reasoning).

[547] Aristotle, in his "*Posterior Analytics"*, is the first source we have that explicitly supports foundationalism through the regress argument, dismissing alternatives like circular reasoning or infinite regress. Then the regress problem became a key focus especially in the writings of ancient skeptics, particularly in Sextus Empiricus's *Outlines of Pyrrhonism* and Diogenes Laertius's *The Life of Pyrrho* in *The Lives and Opinions of Eminent Philosophers*. Eventually, the debate resurfaced in the early 20th century in discussions of scientific methodology (Hasan & Fumerton, 2022). Sellars's (1963) critique of classical foundationalism subsequently sparked discussions that eventually led to defenses of "modest foundationalism" in the 1970s–80s. See the following note for a brief summary of modest foundationalism.

[548] William Alston played a key role in rehabilitating foundationalism by promoting a moderate form of foundationalism, which holds that basic beliefs need only possess some positive epistemic status independent of other beliefs, without requiring infallibility or indubitability. Alston made three important distinctions to address common criticisms: 1) The distinction of epistemic beliefs (e.g., "I am justified in believing X") and non-epistemic beliefs (e.g., "I see a red circle"); non-epistemic beliefs can be justified on their own without needing an epistemic counterpart. 2) The difference between "being justified" and "exhibiting justification"; one can be justified in holding a belief (e.g., Grandma believing she has hands) even if they cannot fully demonstrate or explain that justification, countering the demand for explicit proof for each belief. 3) The separation of immediately justified beliefs (justification regress) from the ability to endlessly demonstrate justification for every belief (showing regress), which critics often conflate. Alston's modest foundationalism allows for basic beliefs that are fallible, revisable, and dubitable but still provide

loop, where claims mutually support one another. As long as the circle is sufficiently broad and interconnected, it can be considered rationally acceptable (Lehrer, 1999). 3) Infinitism: The position that claims must be justified by an infinite chain of non-repeating reasons. By rejecting any stopping point in the justification process, there is no problem of regression (Klein, 1999). Infinitists and coherentists reject foundationalism as arbitrary, claiming that foundational premises lack adequate justification. Foundationalists and infinitists, in turn, criticize coherentism for endorsing circular reasoning, which they argue never truly justifies anything. Lastly, foundationalists and coherentists criticize infinitism as impractical, as humans never possess or can trace an infinite chain of reasoning.[549]

Jusnaturalism, thus, seems to clearly suffers the main foundationalism problem: the challenge of establishing consensus on foundational truths without resorting to arbitrariness. In a paradoxical twist, by attempting to avoid arbitrary laws by grounding them in objective, self-evident moral values, it risks opening itself to subjective interpretations or arbitrary claims about what constitutes "natural" law. For instance, John Finnis, as we have already said in section 3.3.1, identifies seven basic "forms of goods"—life, knowledge, play, aesthetic experience, sociability, practical reasonableness, and religion—which he asserts are not derived from any prior principles but are self-evident and equally fundamental (Finnis, 2011, pp. 85–90). According to Finnis, these goods are universally understood and necessary for human flourishing, in all time and place, but no reasons are given to someone who does not considers (or does not "see") these goods as self-evidently crucial for our flourishing. [550] This is the feature of epistemological foundationalism that is metaphysically problematic for jusnaturalism.

---

sufficient justification for other beliefs (Alston, 1976, p. 180) This view has influenced areas like reformed epistemology, where Alvin Plantinga argues that belief in God can be rational even without direct evidence, as long as it is produced in the right way (Plantinga, 2000). Still, some critics dismiss modest foundationalism for being too weak, arguing that empirical knowledge requires certainty rather than probabilistic justification.

[549] In the next chapter, we will see how constructivism, particularly in its constitutivist version, can serve as an alternative to these models. Here we can briefly anticipate that constitutivism addresses the regress problem by offering a justification that is neither arbitrary nor entirely circular: the constitutive norms stem from rational action itself and are established independently of the overarching network of beliefs. If you deny these norms, you undermine your own standing as an agent. This is the sense in which they are a practical necessity and not arbitrarily chosen as foundations.

[550] Finnis tries to separate his jusnaturalist theory from a strictly theistic foundation, even though religion is one of the basic goods. Nonetheless, Finnis' approach lacks a clear and convincing way to reconcile natural law theory with secular worldviews.

Moreover, the claim that these goods are not derived from any other principles and are equally essential can still lead to disagreements over their interpretation and prioritization, thus introducing concerns about their objectivity. This leads to the other issue connected to the second question raised at the beginning of this subparagraph: the problem of relativity. Philosophical, religious, and cultural disagreements on the nature of morality can lead to skepticism about whether there are some universal sets of natural laws that apply to all people, at all times, and in all contexts. Even in the general idea of flourishing that is behind all the basic good of Finnis, it is possible to argue that there are many competing conceptions of human flourishing that differ across cultures and philosophical traditions (Finnis, 2020). However, disagreement alone does not necessarily undermine jusnaturalist theories or any universalist and objectivist approach. We observe similar disagreements in scientific and historical fields, yet these domains are still treated as objective areas of inquiry. In science, for instance, researchers may debate speculative theories, but such disagreements typically stem from incomplete or conflicting interpretations of empirical data. The underlying assumption is that, with more information, these disagreements can be rationally resolved, gradually bringing us closer to objective truth.[551]

Nevertheless, some relativists or proponents of error theory, such as Mackie, argue that moral disagreements are fundamentally different from scientific ones.[552] Moral disputes, they contend, do not primarily arise from inadequate evidence or incomplete reasoning on some objectives data or standards, as they do in the sciences (Mackie, 1977, pp. 36–37).[553] Instead, they

[551] The fact that a disagreement can be rationally resolved does not mean it will cease to exist. There may still be scientists who reject the standard resolution. However, if the arguments and evidence supporting that resolution are sufficiently robust, dissenters will simply be regarded as outliers. This applies to all forms of disagreement and demonstrates that the persistence of disagreement does not logically imply the absence of objectivity. The fact that some people have denied—and still deny—that the Earth is spherical (or more accurately, geoidal) does not mean the Earth lacks an objective shape. In the same way, moral disagreement does not necessarily entail the nonexistence of objective morality.

[552] One of the key issues in discussions about moral disagreement is that the term often refers to a single phenomenon, while in fact it encompasses a range of related but distinct types of disagreement. These can be, for example: (a) synchronic or diachronic, (b) explicit or implicit, (c) superficial or deep, (d) objective or subjective, and (e) general or particular. Exploring all the nuances of these debates lies beyond the scope of this subsection, which aims only to briefly assess whether—and to what extent—the existence of moral disagreement poses a challenge to jusnaturalism and, more broadly, to any objectivist theory of morality. For further discussion, see Folke Tersman (2022).

[553] Of course, the standards may be an idealization of the way of life from which they arise: the monogamy in which people participate may be less complete, less rigid, than that of which it leads them to approve. This is not to say that moral judgments are purely conventional. Of course, there have been and are moral heretics and moral reformers, people who have turned against the established rules and practices of their own communities for moral

reflect deeply entrenched differences in how people live. In other words, moral disagreements are not simply the result of insufficient evidence; they are rooted in the diverse ways people engage with the world and participate in different life practices of a given society (Mackie, 1977, p. 37).[554] This is supposedly to show that morality is shaped by societal conventions rather than by objective truths.[555]

---

reasons, and often for moral reasons that we would endorse. But this can usually be understood as the extension, in ways which, though new and unconventional, seemed to them to be required for consistency, of rules to which they already adhered as arising out of an existing way of life." (Mackie, 1977, pp. 36–37).

[554] "In short, the argument from relativity has some force simply because the actual variations in the moral codes are more readily explained by the hypothesis that they reflect ways of life than by the hypothesis that they express perceptions, most of them seriously inadequate and badly distorted, of objective values." (Mackie, 1977, p. 37).

[555] Many scholars have recently added empirical arguments to the claim that the prevalence and depth of actual moral disagreement cast doubt on moral objectivism. The most famous example is the case of the Magistrate and the Mob (MM) produced by Peng, Doris, Nichols e Stich (Doris & Plakias, 2008). In this experiment, American participants (mostly of European descent) and Chinese participants residing in the People's Republic of China were asked whether a police chief and a judge did something wrong by conspiring to falsely accuse, convict, and imprison an innocent man. The scenario presented to participants was as follows: 1) A member of an ethnic group committed a terrible crime; 2) A mob, enraged, demanded justice—meaning the identification and punishment of the culprit—otherwise threatening to riot; 3) Unfortunately, the police were unable to identify the actual perpetrator; 4) The riot would have targeted the perpetrator's ethnic group, resulting in deaths and property destruction; 5) The only known way to prevent the riot was to accuse and punish an innocent member of the same ethnic group as the perpetrator—something the police chief and the judge decided to do. Participants were asked questions about moral responsibility, such as whether the potential rioters were morally responsible for the decision to convict the innocent man, and whether the police chief and judge acted immorally by imprisoning him. Several independent studies found that American and Chinese participants disagreed about the moral permissibility of "telishment" (punishing the innocent) in the MM case to prevent an imminent riot. American participants were significantly more likely to believe that the police chief and judge acted immorally and deserved to be punished for their actions. Doris and Plakias (2008, p. 225) argue that MM is a strong candidate for a case of fundamental moral disagreement, because no clear mitigating explanations account for the disagreement. They find it unlikely that either group of participants was less rational than the other, and their responses to relevant non-moral factual questions did not differ significantly. Moreover, they argue that bias is not a plausible explanation because all participants were plausibly disinterested third-party observers when completing the questionnaires. As a result, Doris and Plakias conclude that the usual mitigating explanations do not apply to MM, suggesting it is a case of genuine, fundamental moral disagreement (Doris & Plakias, 2008, p. 230). The absence of references to these types of empirical studies in the metaethical literature does not indicate a supposed arrogance or lack of interest on the part of moral philosophers. Rather, it reflects the fact that the relevance of such findings is often unclear—partly because these studies are not usually designed to address the specific questions that concern philosophers. For instance, they rarely distinguish between different types of moral disagreement, such as whether the disagreement in question is radical or superficial, as noted earlier. Nevertheless, some moral philosophers have directly addressed these issues. Robson, for example, argues that participants may have responded to the scenarios based on "moral intuitions shaped by their cultural moral norms," and may have failed to critically evaluate these norms due to high "cognitive costs" (Robson, 2021, p. 424). From his perspective, neither group's response is necessarily "correct" simply because it prevails within its own culture. Instead, the "correct" response would be the one reached by an ideal, fully informed, and rational agent. This is a form of realist objectivism, which seeks to explain moral disagreement through cognitive flaws, lack of information, different stages of moral development, or differing moral frameworks, with the aim of showing that agreement would be possible under ideal conditions. Constructivists would likely offer a different solution and interpretation of the MM case. Although I cannot delve in detail of the MM

Responding to this challenge entails adopting some form of objectivism. However, moral objectivism is not a monolithic position; rather, it encompasses a variety of theories that differ in how they understand the nature, source, and epistemology of moral truths. For example, the answer that a moral objectivist realist gives is, broadly, the one of Michael Smith when he says that "[m]oral questions have correct answers […] made correct by moral facts." (M. Smith, 1994, p. 9). For realist objectivists, if we ask whether X is moral, we can in principle provide a correct answer that adverts to a fact in the world that pertains to X. Jusnaturalism follows this route, while constructivism follows an objectivist but irrealist path, where the moral inquiry start in "medias res", and we have to explain the "seemingly objective*"* features of our moral experience (Bagnoli, 2002, 2015), but we do so without appealing to mind-independent moral facts (Bagnoli, 2016, p. 1235).

This discussion led to a final tangential issue worth highlighting. Jusnaturalism seems to struggle to adequately reply to the naturalistic fallacy accusation, and more generally to the problem posed by Hume's law, the unwarranted leap from descriptive statements (what is) to prescriptive statements (what ought to be). Clearly, just because something is "natural" in a descriptive sense does not imply it is morally right: aggressive behaviors may be part of human nature (e.g., reacting inappropriately when provoked), but this does not (necessarily) mean that such behaviors are morally permissible. This confusion between what is natural and what is morally right relates to the arbitrariness of jusnaturalism, since moral principles are grounded in nature without clear justification, but merely as self-evident truth.

However, if natural law is understood as what is rational to follow this critique can be addressed. The shift from descriptive to prescriptive is then mediated by reason, not by nature in general, a subtle but important distinction, that makes useful to distinguish between merely jusnaturalist positions, which tends to emphasize the self-evident moral bindingness of the natural property, and *jusratioanlist* position (Villey, 2013), [556] which tends to emphasize the rational way

---

case, since it is an issue that deserves more detailed examination elsewhere, see the next page for a brief anticipation of how constructivism takes distance from realist objectivism.

[556] The distinction, therefore, is not merely etymological but conceptual, and seeks to avoid the ambiguity in the use of the term "nature" not only in the historical passage from Aquinas to modernity, but also among contemporary scholars which at times conceives it as, respectively, metaphysically reified (Finnis, 2011), teleologically revealed (George, 2004) scientifically naturalized (Ruse, 1986), and occasionally rationally constructed (Forst, 1999).

through which we arrive at the normative element. There are, of course, many ways to conceptualize rationality, and some of them may still fall into a form of foundationalism. Neo-Thomistic approaches, for instance, typically belong to this category, as they tend to define reason as the faculty capable of discerning what is objectively good, an external metaphysical property. This perspective, however, does not escape the charge of arbitrariness often associated with Hume's Law. That said, there are alternative positions that successfully avoid this problem. As we saw in section 2.2.3, constructivism does not treat rationality as the capacity to detect or mirror pre-existing moral truths, but rather as a transformative capacity to generate moral principles through the intersubjective act of reasoning. This is how they avoid the dilemma of Hume's law (see 2.3.1) by showing that moral and normative truths are not discovered in the world (as natural or metaphysical facts) but are constructed through rational procedures. We don not infer "ought" from "is" but we construct "ought" from what rational agents would agree to under certain procedural and constitutive conditions. In this way, norms are not derived from facts, but only from the Kantian "fact of Reason", because the source of normativity is practical reason itself (Korsgaard, 1996b).

The problem with jusnaturalism broadly conceived is, to conclude, not that it relies on universally recognizable principles, as relativists or error theorists argue, but rather that these principles are often metaphysical claims about human nature, human purpose (teleology), or the natural world in general, which are reified and treated as self-evident truths. This tendency to treat such intuitions as unquestionable claims reflects the specter of ontological foundationalism and the resulting arbitrariness that haunts jusnaturalism.[557]

[557] Rawls introduces the method of reflective equilibrium as a way to justify moral and political principles without treating them as self-evident truths. Rather than relying on foundational or unquestionable premises, this approach seeks a balance between our considered moral judgments (intuitions about justice) and the principles that best account for and explain them. The aim is to achieve coherence between specific judgments and general principles, ensuring that neither is assumed to be absolutely foundational or fixed, since the reflective equilibrium is a flexible, dynamic process, allowing for continuous revision of both judgments and principles in light of new evidence or more compelling arguments, promoting an evolving and reasoned consistency. While many scholars are dissatisfied with the Rawlsian solution, most of them agree that was one of the first important attempt to propose a viable alternative between foundationalism and formalism. See next chapter, paragraph 5.2.2.

4.2.2 Weakness of Juspositivism: Formalism and Circularity

Juspositivism, as we saw, is usually framed in a negative way. For example, Hart famously identifies legal positivism with "the simple contention that it is no sense a necessary truth that laws reproduce or satisfy certain demands of morality, though in fact they have often done so" (Hart, 1961, pp. 185–186). This definition tells us less about what law is and more about what it does not require: namely, a moral foundation. This is why juspositivism was long associated with the idea that legal philosophy is or should be value neutral, the so-called neutrality thesis. Kelsen, for example, says, "the function of the science of law is not the evaluation of its subject, but its value-free description" (Kelsen, 1967, p. 68) and Hart at one point described his work as "descriptive sociology" (Hart, 1961). But a description of what? As we saw (section 1.1.2 and 1.1.3) scholars like Fuller and Dworkin rightly pointed out that law is dependent not merely on our sensory embodiment but also (as its necessary connections to morality in its most general parts shows) on our moral understanding of moral norms (Dworkin, 1977).

Even if we set aside this explicit relationship between law and morality, many scholars maintain that law cannot be purely descriptive—it is necessarily normative in structure and function, and thus cannot be value-neutral. While some lawyers may treat this claim as a profound insight (and others as a provocation), it is in fact banal. The denial of this thesis (that law could somehow be value neutral) does not even rise to falsity but seems simply to be incoherent. Law is a normative system promoting certain values and repressing others. For example, law is not neutral between victim and murderer or between owner and thief (Green & Adams, 2019)

Constructivists, as we saw, have a way of getting out of this critique by using the distinction between ethics and morality and explaining that the neutrality of law can be reasonably reclaimed if it is an ethical neutrality, but not a moral neutrality. Law is not neutral between victim and murderer because the murderer bears moral responsibility—a form of interpersonal accountability that translates (though does not reduce) into legal responsibility; in this sense, law needs to acknowledge and engage with moral concepts. Yet law can still be ethically neutral in the sense that it does not favor one conception of the good life over another (e.g., a religious life over a secular one); the legal right to refuse life-extending medical treatment reflects an ethically neutral framework in which the state does not impose a specific vision of the good life. Whether one declines treatment based on religious convictions, existential autonomy, or the desire to avoid

suffering, the law does not evaluate these reasons morally—it simply protects the individual's right to make that choice. In doing so, the law permits a plurality of ethical values to coexist, because it assumes that they can be reasonably derived by the same normative assumptions about autonomy, dignity, and human agency.

Of course, juspositivists have not been silent in the face of critiques regarding the supposed value neutrality of law. The neutrality thesis has undergone significant clarification and refinement, especially in the work of figures like Joseph Raz (1970, 1985, p. 199, 1999, 2003a, 2010b).[558] Raz similarly, to constructivists, see Reason and rationality as inherently normative (Raz, 1999, p. 355).[559] More precisely, for Raz rationality is the capacity to recognize and adequately respond to the normative aspects of the world — namely, reasons (Raz, 1999, p. 356)[560] — and he also affirms that this is a constitutive capacity of being a person (Raz, 2010b, p. 21).[561] Practical reasons, specifically, are facts that show what is good in an action and that make the

[558] But even Hart was fully aware that law is used to regulate human conduct in ways that reflect certain social values. His notion of a "descriptive sociology" was not intended to suggest that law lacks normative content, but that a theory of law should first describe the legal system as a social institution before engaging in moral evaluation. Hart's distinction between the internal and external points of view allows him to acknowledge that legal actors operate within a normative framework, even if the theorist describing the system might aim to remain neutral.

[559] "Reason is inherently normative. That is its central characteristic. Therefore, the accounts of normativity and of reason and rationality, though not identical, are inter-related. An account of rationality is an account of the capacity to perceive reasons and to conform to them, and of different forms of conforming to rea sons, and their appropriateness in different contexts. To explain the capacity to conform to reason the account must explain the possibility of error, failure to perceive reasons correctly, and of failure to respond to them once perceived." (Raz, 1999, p. 355). Raz does not distinguish between rationality and Reason, or between being rational or reasonable as constructivists do (Forst, 2012; Rawls, 1971); moreover, the normativity of this capacity is still secondary and derivative from the normativity of reasons, similarly to Scanlon (1998).

[560] "What is capacity-rationality in itself? It is a capacity to see the normative significance of the way things are, to comprehend what reasons they constitute, and the significance of that fact for oneself" (Raz, 1999, p. 356)

[561] "[R]esponsiveness to practical reasons is also constitutive of being a person, for without it there is no action with the intention of doing it. While not all actions are performed for a reason, when we do something with the intention of doing it, which is roughly when we have a purpose in doing it, see a point in doing it, we act for a reason, that is we act in the belief that there is a reason for the action. Such actions, which I will somewhat inaccurately refer to as intentional actions, are governed by reason somewhat less directly than beliefs. Both intentional action and belief are subject to failures correctly to identify reasons, failures to follow through with the implications of reasons one identified (as in failures of memory or lapses of attention) as well as to irrationalities, including motivated irrationalities. In both cases we recognise such failures for what they are in principle, thus acknowledging that beliefs and intentional actions are inherently governed by reason. There is, as we know, a keen debate about the difference between the conditions for rationality of belief and of intentional action, e.g. whether there is epistemic akrasia. For our purpose suffice it to note that whatever differences there are they are secondary to the basic point that we cannot, while acting intentionally, but act for a (perceived) reason (albeit not always the one we believe to be the best reason). Rationality, namely responsiveness to reasons, is thus constitutive of being persons." (Raz, 2010b, p. 21).

action appropriate (Raz, 2009, p. 206).[562] Raz even adopts a conception of morality similar to that of constructivists as the broader normative context (Raz, 1986, p. 398),[563] rather than reducing it to merely psychological or sociological factors, as many other juspositivist scholars tend to do (Leiter, 2007). Moreover, Raz acknowledge an inherent moral function of authority, the so called "service conception of authority", which consists in facilitating conformity to pre-given reason and achieving moral goals that would otherwise be difficult to attain (Raz, 2009, pp. 136–137).[564] This is not however a thesis about what the law is, but rather about what justifies its significance and its authority (Raz, 1985, pp. 298–299).[565]

---

[562] "It seems that rules are reasons for action. A person may well give the fact that his action is required by a rule as his reason for performing it, and an action may be justified because it conforms to a rule. Yet rules are unlike most other reasons. Most reasons are facts which show what is good in an action, which render it eligible: it will give pleasure. It will protect one's health, or earn one money, or improve one's understanding. It will relieve poverty in one's country, or bring peace of mind to a troubled friend, and so on. (Raz, 2009, p. 206)

[563] "This use of 'moral' is compatible with the one described in Chapter Twelve. There we noted that from the perspective of an agent certain reasons are not normally regarded as moral. Here we note that when considering practical reason generally, without being in the position of any agent in particular, all values and reasons are moral." (Raz, 1986, p. 398).

[564] "The suggestion of the service conception is that the moral question is answered when two conditions are met, and regarding matters with respect to which they are met: First, that the subject would better conform to reasons that apply to him anyway (that is, to reasons other than the directives of the authority) if he intends to be guided by the authority's directives than if he does not (I will refer to it as the nor mal justification thesis or condition). Second, that the matters regarding which the first condition is met are such that with respect to them it is better to conform to reason than to decide for oneself, unaided by authority (I will refer to it as the independence condition). Simple examples of regulations regarding dangerous activities or materials illustrate the point. I can best avoid endangering myself and others by conforming to the law regarding the dispensation and use of pharmaceutical products. I can rely on the experts whose advice it reflects to know what is dangerous in these matters better than I can judge for myself, a fact that is reinforced by my reliance on other people's conformity to the law, which enables me to act with safety in ways that otherwise I could not. Of course, none of this is necessarily so. The law may reflect the interests of pharmaceutical companies, and not those of consumers. If that is so it may lack authority over me because it fails to meet the normal justification condition. But if it does meet the normal justification condition it is likely to meet the independence condition as well. Decisions about the safety of pharmaceutical products are not the sort of personal decisions regarding which I should decide for myself rather than follow authority." (Raz, 2009, pp. 136–137). The service conception, in short, shows how an authority (including legal authority) can be legitimate and justified on moral reasons: it states that an authority is legitimate only if the subjects to whom it is addressed are likely to better conform to the reasons that already apply to them (reasons other than the purported authoritative directives) by accepting the directives of the supposed authority as authoritatively binding and trying to follow them, rather than by attempting to follow those reasons directly on their own. Authority, then, is justified if it helps people do better what they already ought to do (the pre-existing reasons they already had).

[565] To be more precise, the service conception of authority includes three theses. 1) *Dependence Thesis* (Authoritative directives should be based on reasons that already apply to the people subject to the authority; these are called dependent reasons—reasons that pre-exist and justify the directive). 2) *Normal Justification Thesis* (an authority is justified if following its directives helps people better conform to the reasons that already apply to them, more so than if they acted on their own). 3) *Preemption Thesis* (when an authority issues a directive, its command replaces some of the reasons the subject would otherwise consider, rather than being added to them. This avoids double-counting reasons). Together, these three theses explain why acceptance of authority isn't blind obedience and that it

By contrast, Raz's crucial thesis about the nature of law is what he calls the "Sources Thesis", which holds that the existence and content of law can be determined solely on the basis of social facts, without requiring reference to moral reasoning (Raz, 1979, p. 53, 1985, p. 296). A legal system (e.g., a specific statute or precedent) is by itself a social fact, "posited" through human activities such as legislation, custom, or judicial decisions. Moreover, to know what constitutes a law within a given legal system, it is not necessary to assess its moral merit, because law operates for citizens at the executive stage of the social decision-making process, not the deliberation process (Raz, 1970, p. 214).[566] This means that its directives are designed to establish reasons for action in an authoritative and final manner, without requiring individuals to re-deliberate the underlying moral reasons. The aim is "finality", not necessarily certainty or predictability (Raz, 1970, p. 216).[567]

Still, how can the law's validity be entirely determined by social sources, as the Sources Thesis claims, while its authority is justified only if it reflects and helps people better conform to

---

must be justified by the authority's ability to improve the subject's reasoning and action in a given context. "The first two theses articulate what I shall call the service conception of authority. They regard authorities as mediating between people and the right reasons which apply to them, so that the authority judges and pronounces what they ought to do according to right reason. The people on their part take their cue from the authority whose pronouncements replace for them the force of the dependent reasons. This last implication of the ser vice conception is made explicit in the preemption thesis. The mediating role of authority cannot be carried out if its subjects do not guide their actions by its instructions instead of by the reasons on which they are supposed to de pend. No blind obedience to authority is here implied. Acceptance of authority has to be justified, and this normally means meeting the conditions set in the justification thesis." (Raz, 1985, p. 299).

[566] "Law exists, according to the sources thesis, only in societies in which there are judicial institutions which recognize the distinction between the deliberative and executive stage, that is they hold themselves bound to recognize and enforce certain reasons not because they would have approved of them had they been entrusted with the question in the deliberative stage but because they regard their validity as authoritatively settled by custom, legislation, or previous judicial decisions so that the matter is held by the courts in the litigation in front of them as being in the executive stage. When this is the case the courts will not entertain moral arguments about the desirability of regarding a certain fact (e.g. a previous enactment) as a reason for a certain action but will once the existence of the relevant fact has been established through morally-neutral argument hold it to be a reason which they are bound to apply. Only reasons binding on the courts in such a way, i.e. only 'executive' reasons, reasons the existence of which can be established without invoking moral arguments, are legal reasons. The sources thesis assigns the law to the executive stage of social decision-making." (Raz, 1970, p. 214).

[567] "Finality is a function of a division of labour between formal and informal stages of reflection and deliberation and formal and informal stages of implementation. Even in the life of a person deliberation is often followed by a decision which may anticipate action by a considerable period of time and which fixes the intention which thus becomes relatively immune from revision. Where social action is concerned, the need to fix the decision in advance of the action, the need to decide authoritatively what is to be done in a way which is binding on members of the community so that they are not allowed to differ from each other because of disagreement on what is best--such need for finality becomes overwhelming." (Raz, 1970, p. 216).

moral or rational reasons, as the Service Conception suggests? On the surface, the two theses appear to pull in opposite directions—the former grounding law in morally neutral social facts, the latter grounding legitimate authority in moral reason. Furthermore, Raz also suggests that the moral reasons that makes the authority of law legitimate according to the service conception—particularly the preemption thesis—are not merely reasons for action, but *exclusionary* reasons (Raz, 2009, p. 144).[568] Preemption clarifies that authority does not simply add another reason to the decision-making balance; rather, it replaces the subject's own weighing of reasons with the authority's judgment. In other words, an exclusionary reason is a second-order reason—one that excludes the subject from acting on conflicting first-order reasons that the authority is presumed to have already considered. If the law is legitimate, it shapes and specifies the application of morality by acting as a second-order reason that overrides certain first-order moral considerations, which it is meant to reflect and incorporate (Raz, 2009, p. 216).[569]

Raz clarifies that in order to be capable of claiming authority, a system of directives (such as law) must meet two non-moral conditions (Raz, 1985, p. 303).[570] 1) It must be, or at least present itself as, the judgment of someone regarding how the subjects ought to behave. This reflects the

[568] "A proper understanding of preemption removes any suspicion of a problem. A binding authoritative directive is not only a reason for behaving as it directs, but also an exclusionary reason, that is a reason for not following (ie not acting for) reasons that conflict with the rule. That is how authoritative directives pre-empt. They exclude reliance on conflicting reasons, not all conflicting reasons, but those that the law-maker was meant to consider before issuing the directive. These exclusionary reasons do not, of course, exclude relying on reasons for behaving in the same way as the directive requires. Think about it: authority improves our conformity with reason by overriding what we would do without it, when doing so would not conform with reason. So, assuming that it is entirely successful in its task, it need not and does not stop us from following the reasons on the winning side of an argument. It must, however, if it is to improve our conformity with reason, override our inclination to follow reasons on the losing side of the argument. Hence the pre-emption excludes only reasons that conflict with the authority's directive." (Raz, 2009, p. 144).

[569] "I suggested that rules, rules of the kind we are considering, are not simple reasons but a structure of interrelated reasons. They are, first, reasons for the acts they prescribe, but they are also, second, reasons not to act for some competing reasons. The rule that no more than three guests may be invited by a club member is, first, a reason for members not to invite more than three guests, and also, second, a reason not to act on some reasons for inviting a fourth guest. I call the second kind of reasons exclusionary reasons (for they exclude action for some reasons) and I call the rule itself a protected reason, because the reason for the action it prescribes is protected by these exclusionary reasons." (Raz, 2009, p. 216).

[570] "If this view is correct then since the law necessarily claims authority, and therefore typically has the capacity to be authoritative, it follows that it typically has al the non-moral, or non-normative, attributes of authority. The remainder of my argument, however, does not depend on this strong conclusion. We will concentrate on two features which must be possessed by anything capable of being authoritatively binding. These two features will then be used to support the sources thesis." (Raz, 1985, p. 303).

mediating role of authority: it operates on reasons that already apply to us, based on the presumption that following it will help us better conform to those reasons. 2) It must be identifiable as issued by the presumed authority without relying on the very reasons or considerations that the directive itself is meant to settle.

It is precisely this second condition that connects the Sources Thesis with the Service Conception: for an authoritative directive to fulfill its mediating and preemptive role—that is, to provide a *settled* reason that replaces the need for direct deliberation on the underlying reasons—its content and existence must be identifiable independently of those underlying reasons. If identifying the content of a law required first engaging in complex moral reasoning to determine what the "right" or "best" reason was, law would fail in its function of simplifying and settling matters. Its very purpose is to provide clear and definitive guidance that relieves subjects from having to independently recalculate the balance of reasons. Therefore, the Sources Thesis, by holding that the identification of law must rely on social facts (such as legislative enactment or judicial custom) rather than on moral argumentation, is not a contradiction, but a precondition for the intelligibility of law's claim to authority (Raz, 1985, p. 307).

This leads to the classic objection raised against juspositivists: the problem of unjust laws. Raz acknowledges that legal systems can possess systemic moral properties (the most important one is again the ability to ensure social coordination and public order, the absence of which would make the pursuit of many values impossible), even though individual laws within such systems may be unjust or morally questionable (Raz, 2003a, pp. 6–7). The Sources Thesis merely explains how law can be the kind of thing that claims authority, while leaving open the question of whether that claim is actually justified in a given context. Moral arguments can be part of the substantive content in the act of legislation, but they are not part of the criteria for legal validity. Judges, being human beings, are nevertheless subject to morality (Raz, 2009, pp. 184–190)[571], and may

[571] "This, then, is finally the full answer to the question of why judges are humans too, why they are subject to morality: they would not be subject to the law were they not subject to morality. Let this reply not be misinterpreted: it does not deny that we can discuss and describe the law from a detached point of view. We can talk of a legal system that is practised in a country and of its requirements and implications without making any assumption that they have normative standing.[13] Such discourse, free of the assumption of the normative standing of the law, is the equivalent of talking of what people demand of others without implying that these demands have any normative standing, that the others have reason to comply with them as they are intended to do. But—and this is my point—where the law is normatively valid, it is so in virtue of a moral principle, and therefore if we take the law to be normatively valid we cannot construe its requirements as constituting a point of view independent of morality, a point of view that represents a separate normative concern that has nothing to do with morality, and then ask whether it recognizes morality as

sometimes implement it, but this does not mean that laws must be subordinated to morality, by conceiving law as a mere implementation of morality. If a law requires the application of moral considerations (e.g., "a contract is valid only if it is not immoral"), for Raz, the law itself is identified through its source (the legislator who enacted it), but the morality it refers to is not thereby "incorporated" into the law in the strict sense. Its function is rather to *modulate the exclusion* of morality, indicating to judges that, in this particular case, moral considerations are not excluded from their decision-making.

Raz uses an analogy with private international law to clarify its position: when the law of one country refers to the application of a foreign law, that foreign law does not thereby become part of the domestic legal system; similarly, a law that refers to moral considerations does not make those considerations part of the domestic law (Raz, 2009, p. 201).[572] The distinction between what is part of the law and what is applicable according to the law but not part of it is fundamental to understanding the unique nature of the legal system for Raz. A law is legally valid if it is enacted according to the rules and procedures recognized by a legal system—such as being passed by a legitimate legislature—regardless of whether it is morally good or bad. Therefore, a law is valid if it was created through the proper procedures recognized by a legal system (like passing through parliament), *regardless* of whether it's morally right or wrong.

In sum, Raz does not claim that morality plays no role in law, as some have interpreted him—both among his critics and his admirers (Shiner, 1992). On the contrary, he conceives law as a social institution with a profound moral function: to serve the reasons individuals already have, and to coordinate social action. However, in order to function as an authoritative system that resolves disagreement and provides public standards, law must be identifiable independently of

applying to its officials. The boot is on the other foot; the question is whether morality, which applies to all humans simply because they are humans, has room for the law." (Raz, 2009, p. 189).

[572] "The conclusions of the earlier part of this chapter pose a difficulty for the incorporation thesis. If morality applies to people and courts alike anyway, then we are all, courts included, bound by it even before its incorporation. In what way can incorporation turn it into law? Th e fact that the incorporation thesis is the purest expression of the simple view counts against it rather than being a point in its favour. We know that the simple view is too simple. We know that Polish law is not part of Greek law just because Greek conflict-of-law rules direct people and courts to follow Polish law on certain occasions. T h e incorporation thesis claims that moral standards turn into law simply because of their incorporation. It seems to lack the resources to distinguish between law directing us and the courts to follow some foreign law or to obey the rules of some associations, and so on, and the incorporation of morality." (Raz, 2009, p. 201).

any moral analysis of its content. The "morality" Raz excludes from the identification of law is the evaluative judgment of its merit or justice — not morality understood as the set of practical reasons that may justify law's existence and its claim to authority. Thus, law is not moral *in its content* (Sources Thesis), but it does make a moral *claim* in its function and in its assertion of legitimacy (Service Conception of Authority).

Constructivists are likely to sympathize with many of the more moderate reinterpretations of juspositivism like that of Raz, particularly those that acknowledge the normative dimensions implicit in law's practical function; what they would still sharply contest is the affirmation of the Sources Thesis, and in particular, Raz's claim that law operates for citizens at the executive stage of the social decision-making process, rather than at the deliberative stage (Raz 1970, 214). This view, by design, separates legal authority from another crucial context of normativity, that of the political, in which citizens actively reason, deliberate, and contest norms. Forst, as we saw, stressed vigorously that citizens are both authors and addressees of the law: they are not merely recipients of pre-established directives but are fundamentally involved in the processes through which legal norms are constructed and justified (Forst, 2002, pp. 127–128, 2024, p. 13).

This divergence is also crucial to properly understand the most enduring critiques of legal positivism: its perceived inability to adequately respond to the problem of unjust laws. Although Raz attempts to account for this issue, as we saw, by distinguishing between the identification of law and its moral justification, his framework still struggles to fully address the deeper concern: how to resist or delegitimize legal systems when those in power are morally corrupt or systematically unjust. The positivist model's emphasis on obedience to validly sourced rules, irrespective of their moral content, leaves it vulnerable to the charge of normative powerlessness—particularly in contexts where law is wielded as a tool of domination rather than justice (Forst, 2017a, pp. 154–159).

This critique is typically framed within the lens of formalism, which, in its broadest sense, emphasizes form over content or meaning (Frege, 1903). Juspositivists are often accused of formalism because they treat law as a self-contained system whose legitimacy rests on its formal attributes: clarity, internal coherence, and conformity to established procedures. As Frederick Schauer explains, formalism is a decision-making style that seeks to constrain the range of reasons deemed legitimate in legal reasoning—often justified in the name of objectivity, predictability, or

neutrality (Schauer, 1988). However, this procedural focus becomes problematic when it assumes that legality—and by extension justice—can be fully realized through procedural correctness alone, regardless of the moral quality of the law's content. The main fear is that substantive justice, under this view, becomes irrelevant to the validity of legal norms.[573]

A direct consequence of this formalist orientation is a narrowed conception of the rule of law. Rather than viewing the rule of law as an overarching principle by which a legal system should be morally assessed, it is reduced to a set of procedural criteria internal to the legal system itself. This divorce of the rule of law from substantive conception of justice highlights that the formal approach is purely "instrumental" (George, 2001, p. 251). Raz famously likened the rule of law to a knife, which of course had no moral value, but could be used effectively for both good and bad purposes (Raz, 1979, p. 225); hence Raz's provocative statement that "gross violations of human rights" are compatible with the rule of law (Raz, 1979, p. 213, 2010a).

Raz's position becomes especially problematic when he invokes morally resonant terms such as "freedom" and "dignity" (Raz, 1979, p. 220, 1986, p. 75), but redefines them in a purely formal and procedural way derived from conformity to formal procedures, even when the content of the law may contradict the substantive values those terms typically express. Yet, this raises a troubling paradox: how can a legal system meaningfully claim to respect "human dignity" while simultaneously permitting institutionalized injustices or gross violations of human rights? For example, it seems both perplexing and contradictory for a legal system to claim that a country practicing institutionalized child slavery respects human dignity simply because its servitude laws are prospective, clear, publicly accessible, and relatively stable—while another country that robustly protects human rights is deemed deficient merely because it has a complex and technical tax code.

This tension echoes a broader issue: the problem of arbitrariness that, as we said in the previous section, is a challenge that both jusnaturalism and juspositivism grapple with, albeit in different forms. On one side, if the foundation of law rests on substantive moral truths, these may appear arbitrary to those who do not accept them as self-evident or universally compelling (they offer no further justification to the unconvinced); on the other side, if law is grounded purely in

[573] Inclusive versions of juspositivism will try to avoid this, as we mentioned in section 4.1.4 (Coleman, 2003).

the institutionalization of social facts, then arbitrariness arises from the fact that these facts lack intrinsic normative authority (they are binding simply because they are socially accepted, not because they are morally justified). Paradoxically, juspositivism has been accused of both. For example, Trevor Allan (2005, pp. 62–67) argues that Raz's approach is actually 'based upon substantive foundations" beside is professed Sources Thesis. At the same time, juspositivists are frequently challenged for embracing some sort of social arbitrariness that render legal authority vulnerable to the whims of cultural habit (Dworkin, 1977, pp. 42, 248–260, 2011, pp. 166–167) or, even worst, of power structures (Forst, 2017a, pp. 38–44, 2024, pp. 155–157, 163–167).

Specifically, these critics usually formulate the critique of arbitrariness in terms of *circularity* (Dworkin, 2011, p. 404),[574] where juspositivists conceive the legal system as validating its own norms by appealing to procedures internal to the legal system itself—for instance, by asserting that a rule is valid because it conforms to the "rule of recognition," which is valid simply because it is accepted by officials. This creates a potentially vicious circle, in which the legal system becomes self-validating and impervious to external critique. The problem, therefore, is not just that positivist accounts can validate unjust laws; it is that they cannot convincingly explain why anyone should treat those laws as authoritative in the first place, especially in cases where the content of the law conflicts with fundamental moral norms.

However, this does not mean that all forms of proceduralism lead to formalism or vicious circularity. Many scholars stressed that legal procedures can be normatively robust if they are capable of self-correction, responsiveness, and open to contestation. For example, Roberto Unger, a leading figure in the Critical Legal Studies movement (a movement that criticized the formalist drift of juspositivism) advocated for a "destabilized" conception of law, in which legal rules are continually reinterpreted in light of shifting political and ethical perspectives. From this view, the problem does not lie in proceduralism itself, but in a formalistic understanding of procedure—one that treats legal processes as neutral, closed, and disconnected from normative deliberation (Unger,

[574] "We cannot solve the circularity problem of the two- systems picture through an analysis of the concept of law unless that concept can sensibly be treated as a criterial (or perhaps as a natural- kind) concept. But it cannot be. There is no agreement among lawyers and judges in complex and mature political communities about how to decide which propositions of law are true. No wonder positivists have had such difficulty in explaining the kind or mode of conceptual analysis they have in mind." (Dworkin, 2011, p. 404).

2015, pp. 78–80).[575] Legal positivists must therefore engage seriously with this line of critique if they wish to preserve the normative legitimacy of law. Merely refining internal procedures or appealing to systemic coherence will not suffice if those procedures are perceived as indifferent to questions of justice or democratic participation.[576] In order to avoid the pitfalls of circularity and formalism, juspositivist theories need to articulate how legal procedures can remain open to contestation and responsive to the values and claims of those subjects to the law.[577]

4.3 MOVING BEYOND THE RESULTING THEORETICAL STALEMATE THROUGH UNIVERSALITY

As was he have shown, the long-standing divide between jusnaturalism and juspositivism is frequently cast as a contrast between metaphysical foundationalism and procedural formalism. This debate does not clearly favor one side over the other, and becomes more complex upon closer examination, particularly when one considers the internal tensions and conceptual difficulties inherent in each approach, like arbitrariness and circularity that, as our analysis has made clear, are, although counterintuitively, not unique to either foundationalism or formalism, but can be directed, in a qualified sense, at both, revealing deeper structural limitations shared across the divide.

Foundationalism risks arbitrariness when it asserts moral truths as self-evident or objective starting points, without offering any shared or intersubjective basis for their selection; formalism, although it presents itself as avoiding such arbitrariness by grounding normativity in the

[575] Unger defines formalism precisely as the belief that lawmaking differs fundamentally from law application (Unger, 2015, pp. 78–80).

[576] Which are not separated political claims, but crucial normative elements that intersect across the practical legal context and the practical political context.

[577] This is precisely what constructivists have tried to do, by developing a conception of legal process that is both more grounded and more open, by embracing procedural reflexivity as a means of ensuring legitimacy. As Habermas said, procedures are not valuable merely because they are followed, but because they embody deliberative legitimacy (Habermas, 1996). The idea is that if proceduralism is reflexive instead of merely formalistic, the circularity of the process is not vicious, but a virtuous one.

procedures and internal coherence of a system, might be accused of arbitrariness in the selection phase of the foundational procedure or rule from which legal or moral validity is to be derived.[578] Moreover, foundationalism might be accused of circularity when the truth of the foundational principle is defended by appealing to conclusions that already depend on its assumed validity (thus presupposing as self-evident what it was meant to ground); formalism, similarly, is circular in that it treats the validity of norms as derived from their conformity to procedural rules, while the authority of those procedures is accepted simply because they are part of the system (the rules gain legitimacy from the system, and the system gains legitimacy from the rules, with no external criterion offered to break the circle).

More generally, what unites these two modes of reasoning is a deeper failure to provide a publicly accessible and mutually justifiable basis for legal authority. Jusnaturalism, in assuming the self-evidence or unquestioned acceptance of its grounding principle, is unable to persuade those outside its conceptual or moral framework.[579] Juspositivism, in treating procedural conformity as both necessary and sufficient for legal validity, is equally unable to justify the system itself to those who question its foundational authority. In both cases, the law is insulated from critical evaluation by the very structure that purports to justify it. The consequence is a form of 'normative closure' that sits uneasily with the pluralistic and contested nature of democratic societies, where legitimacy depends not merely on formal validity or metaphysical principles, but

[578] For example, the rule of recognition that Hart places at the beginning of its legal framework as the meta-rule that gives coherence to it, emerges from accepted practice and carries internal logic, but it is not justified per se. That is, there are no independent moral or normative reasons to prefer one specific rule of recognition over another. This does not imply that law is wholly arbitrary: most legal systems incorporate internal mechanisms—such as principles of fairness, equality before the law, and procedural rights—that serve to constrain arbitrariness and support the legitimacy of legal outcomes within the system; yet the arbitrariness of the initial grounding remain a valid concern

[579] Jusnaturalists like Finnis might argue that no one can stand outside a minimal moral framework grounded in basic goods that they cannot but share (Finnis, 2011). While this may resemble the constitutivist move often found in constructivist theories, it ultimately differs in important ways. The key issue is that the jusnaturalist framework is not merely *constitutive* but also *substantive* in a metaphysically problematic sense. Unlike constructivism, which seeks to ground normativity in the structure of practical reason itself, jusnaturalism tends to assert the self-evidence of specific moral truths or goods, often without offering justification accessible to those who do not already accept them. Although debates around the inescapability of practical reason remain ongoing and far from settled (Enoch, 2006; Ferrero, 2009), a theory that posits practical reason as constitutive, without relying on ethically substantive goods such as religion, is better positioned to respond to critiques that challenge its universality. Constructivism aims to demonstrate that even those who appear to stand outside its moral or conceptual framework are, in fact, already engaged in normative reasoning, but its argument is also not only internal, but also public: it does so by employing a non-substantive conception of morality—one that is not grounded in particular ethical or metaphysical values, but instead in principles of public justification and reciprocal reasoning (Bagnoli, 2016; Forst, 2017a).

on the capacity of legal systems to offer reasons that can be accepted by free and equal citizens regardless of their substantive ethical commitments.

The fact that these two positions cannot escape these problems if rigidly conceived as above is something that has been repeatedly emphasized by constructivists. As Habermas points out in a rereading of his book "*Faktizität und Geltung*": "Like legal realists, who treat legal norms just as the result of policy-decisions, positivists cannot explain how legitimacy can spring from sheer legality. Both positivists and realists refuse to recognize any claim to legitimacy stronger than the kind of legal validity that terminates in formally correct enactment and efficient enforcement. Proponents of natural right theories, on the other side, derive the legitimacy of positive law immediately from a higher moral law […] Even if we leave problems of foundationalism aside, such an assimilation of law to morality blurs important differences between the two. Whereas moral norms primarily tell us, what we ought to do and what we owe each other, modem law is in the first place designed for the distribution of individual liberties-for the determination of private spheres where everybody is free to do what one wants to do. Moral rights, on the other hand, are derivative from other people's duties towards us, whereas in law, rights are duties, since legal duties only result from mutual constraints of equally granted liberties." (Habermas, 1999, p. 938).[580]

Therefore, according to Habermas, law should neither be entirely separated from morality, nor should the legitimacy of law be entirely absorbed by legal validity (juspostivist thesis); at the same time, the universality of law should not be conflated with morality, as this would obscure the distinct types of legitimacy inherent to each normative sphere (jusnatural thesis). This normative element, it is said, it is not fixed as the physical and can vary widely depending on cultural, religious, and philosophical perspectives. This is the challenge of the moral disagreement for jusnaturalism: how we can assert universal recognition when different societies may have conflicting moral beliefs and practices? Rigid jusnaturalism can only answer this problem by

[580] Habermas discusses legal positivism and legal realism together, noting that both schools struggle to justify how legitimacy arises solely from legal validity, as they tend to view legal norms merely as products of policy decisions. However, to provide an accurate historical context, it is important to note that legal realism emerged after legal positivism and sought to move beyond it. While legal positivism established the foundation for understanding law as a system of rules and institutions distinct from morality, legal realism critiqued the formalism of positivism by emphasizing the role of social, economic, and political factors in shaping legal outcomes. Both perspectives continue to influence contemporary legal theory and practice, contributing to ongoing debates about the nature and legitimacy of law.

redirecting the source of universality in some other entities, in order to explain the moral disagreement only as factual disagreements. What is lacking in a jusnaturalist approach is, therefore, a definition of universality that is not based on a foundationalist metaphysic. Conversely, the main problem for a strict positivist approach is the negation that morality and laws are connected: a law is valid as long as it is duly enacted and recognized within a legal system, regardless of its moral content. This stance can lead to situations where we have no theoretical tool to critique laws that are considered morally repugnant or evidently unjust like in Third Reich case. What is lacking in a juspositivist approach is, therefore, a definition of universality that is not wholly contained inside legal validity.

To move beyond the respective limitations of both jusnaturalism and juspositivism, what is required is a constructivist account of universality—one that neither relies on metaphysically substantive assumptions nor retreats into a procedural empty formalism. Constructivism offers a middle path by grounding universality in the shared structures of practical reasoning, public justification, and reciprocal deliberation. This is the task we will take up in the next chapter: to develop a constructivist conception of universality capable of accounting for the universal authority of human rights without falling into the foundationalist or formalist traps of earlier theories.

CHAPTER 5 Grounding Universality in Constructivist Terms

To look for a single general theory of how to decide the right thing to do is like looking for a single theory of how to decide what to believe (Nagel, 1979, p. 135).

The idea of universality has long been central to philosophical inquiry. Unlike mere generality—which describes how widely a concept or phenomenon happens to apply—universality concerns the form of validity that a claim, concept, or norm possesses; to call something universal is not simply to say that it holds true in many or all cases, but that it *ought* to hold for any rational subject, regardless of specific conditions. This is a central concern in many fields, ranging from metaphysics and epistemology to aesthetics and linguistics. In moral and political philosophy specifically, universality is a crucial concern: it seeks to establish principles that apply across cultures, times, and individuals—offering a response to moral particularism and relativism, as well as to political fragmentation and polarization. At stake is the question of whether there exist moral and political norms that transcend historical contingency, local custom, or personal preference.

As we saw in the introduction, the universality of morality is one of the key philosophical assumption of Kantian constructivism, that distinguishes it from other moral theories (Williams, 1985; Joyce, 2001) and also from other versions of constructivism in a qualified sense (Street, 2008). It is beyond the scope of this project to persuade those who remain unconvinced by the appeal of the universality assumption. However, I will briefly outline three practical reasons why adopting it remains both reasonable and advantageous: 1) Grounding the possibility of Reasoning and Deliberation. Universality is not only a moral assumption but a logical precondition for any

form of reasoning and deliberation, that aims at justification rather than mere agreement or negotiation. Without it, judgments and arguments would lose coherence: if norms or reasons applied only to some and not to others, there would be no shared ground for persuasion or critique, and deliberation would collapse into relativism. Universality thus provides the structural condition within which rational justification and critique become possible, without eliminating the space for reasonable disagreement or plural perspectives. 2) Guiding principle for fairness. Universality also operates as a regulative principle in the construction and application of norms. By excluding selective exemptions and privileging no particular standpoint, it protects practical reasoning from arbitrariness and bias. This equal applicability reinforces fairness and secures the legitimacy of obligations and judgments across different agents and contexts. 3) Accountability beyond boundaries. Universality establishes standards by which individuals, institutions, and states can be held responsible, even when local laws or cultural practices conflict with broader principles of justice. By subjecting all actors to equal scrutiny, universality strengthens democratic global justification.

After all, while universality in morality and politics may appear as a deeply contested notion, it is, in another sense, almost trivial. Even an error theorist like J.L. Mackie, who denies the objective truth of moral claims, admits that "moral judgments are universalizable" and that "this principle, in some sense, is beyond dispute" (Mackie, 1977, p. 83). A similar position is found among several political theorists and scholars of human rights who, while critical of strong metaphysical universality, nonetheless affirm the universalizability of human rights claims as a feature of their normative grammar (Donnelly, 2002; Ignatieff, 2003). The deeper philosophical question, then, is not whether some form of universality is present, but how it ought to be properly characterized. What kind of universality is at stake? Is it formal or foundational, procedural or substantive? Does it entail a shared human nature, rational consensus, or simply a logical consistency and coherence in practical reasoning?

In this final chapter, we will examine the constructivist conception of universality. We begin by briefly reconstructing how the notion was conceptualized by Kant, appropriated and elaborated by Rawls and Habermas (section 5.1) and subsequently refined by more recent constructivist thinkers (section 5.2). We will then explore how this constructivist account of universality avoids the pitfalls of formalism and foundationalism, as well as the problems of

arbitrariness and circularity, which we previously identified in both jusnaturalist and juspositivist traditions (section 5.3). In this way, we will be able to verify whether the constructivist conception can be regarded as an innovative contribution to the field.

5.1 EARLY CONSTRUCTIVIST CONCEPTIONS OF UNIVERSALITY

Contemporary constructivism emerged, as we explained along this project, as a renewed interpretation of Kant aimed at offering an alternative to both traditional deontological ethics and formalist readings of moral theory (Herman, 1993, pp. 208–243; O'Neill, 1989, pp. 201–233). As we showed (section 2.2.4), in this constructivist reading Kant's notion of universality is reinterpreted not as a rigid, top-down imposition of rules but as a procedural feature of practical reason for understanding the structure and source of normativity: the procedure of universalization—asking whether a maxim could be willed as a universal law—is seen as a way of constructing normativity from the standpoint of practical reason itself, rather than discovering it in some external moral reality (Reath, 2013, p. 41).[581] Kant's notion of universality is therefore not primarily normative, but meta-normative: the central concern is not to deliver a substantive list of duties, but to use the universalization procedure to explain why and how normative obligations can be justified in the first place (G 4:461). The Categorical Imperative is, in this sense, a meta-normative tool to explain the source and authority of normativity.

Moreover, constructivists, by emphasizing the deep connections between the three formulations of the Categorical Imperative—the Formula of Universal Law, the Formula of Humanity, and the Formula of the Kingdom of Ends—have shown that Kant's notion of universality is crucially intertwined with a public conception of reason (Sensen, 2011, p. 109). To universalize a principle is not merely to apply it consistently to oneself, but to imagine it as a principle that could be publicly endorsed, accepted, and acted upon by all rational agents under similar circumstances. In this sense, moral reasoning is not a solitary or private activity, but a form of public justification, where one must be able to offer reasons that others could recognize,

[581] The will wills itself only insofar as it is "self-consciously guided by […] the formal end of satisfying its own internal norm (the conditions of universality)" (Reath, 2013, p. 41)

understand, and, in principle, accept. A maxim is thus universally valid only if it could be willed as a law by everyone without contradiction—not just in terms of logical contradictions (contradictions in conception), but also of practical contradictions (contradictions in the will). This reflects Kant's broader vision of the Kingdom of Ends, in which each individual is both a legislator and a subject of moral law—recognized as a free and equal rational being, capable of self-legislation and of participating in the collective construction of moral norms that are universally and mutually binding (G 4:426).

This intertwining of universality and publicity represents a crucial legacy that allows us to regard both Rawls and Habermas as Kantians,[582] despite the fact that each accused the other of misrepresenting the very idea of public reason (Rawls, 1995, pp. 132–133; Habermas, 1995, pp. 114–117). To unpack this tension, we will begin with Rawls's conception of public reason, which offers a useful entry point for examining the role of universality within his broader philosophical framework.

5.1.1 Rawls's Original Position and Reflective Equilibrium

Rawls introduced the concept of public reason as a response to the "fact of pluralism," where citizens adhere to diverse and often incompatible comprehensive doctrines (religious, philosophical, moral), often incompatible with one another (Rawls, 1997, p. 776).[583] Within this

[582] For Finlayson they should be called Post-Kantians, "in the sense that they start from Kantian assumptions and that they welcome a Kantian interpretation of their enterprise, while rejecting the metaphysical assumptions behind Kant's project, such as his two-worlds metaphysics and the strict transcendental–empirical distinction" (Finlayson, 2019, pp. 5–6). This label, however, seems to conflate two notions of metaphysics, assigning to Kant a notion that he explicitly tried to combat in the first Critique, namely a metaphysics that tries to give answers beyond the realm of possible experience; the step required, instead, is merely into a "metaphysics of morals" (G 4:427). What Kant means by this is a "system of a priori cognition from concepts alone" (RL 6:216; G 4:412).

[583] Some scholars emphasize that, beside this declaration of intent, the values assumed at the outset of Rawls's framework are nonetheless inherently liberal-democratic; this has led to the enduring critique that Rawls never fully escapes a reliance on the very political culture he seeks to justify (Sandel, 1982). Yet, Rawls does not attempt to avoid this charge: he openly embraces it. As he explicitly states: "The idea of public reason specifies at the deepest level the basic moral and political values that are to determine a constitutional democratic government's relation to its citizens and their relation to one another. In short, it concerns how the political relation is to be understood. Those who reject constitutional democracy with its criterion of reciprocity will of course reject the very idea of public reason. For them the political relation may be that of friend or foe, to those of a particular religious or secular community or those who are not; or it may be a relentless struggle to win the world for the whole truth. Political liberalism does not engage those who think this way." (Rawls, 1997, pp. 766–767). Rawls' primary purpose is to ensure that the exercise of political power is legitimate and can be reasonably justified "in accordance with a constitution the essentials of which

practical context, public reason describe the mode of reasoning appropriate for citizens deliberating in the public political forum about constitutional essentials and questions of basic justice (Rawls, 1993, p. xxi).[584]

A citizen engages in public reason when they "deliberate within a framework of what they sincerely regard as the most reasonable political conception of justice." (Rawls, 1999, p. 140). This practice must produce reasons and values that other citizens, as free and equal individuals, could reasonably be expected to endorse, and it represents the way in which the citizens of a constitutional democracy constitute the basic structure of a bounded society (Rawls, 1999, p. 136).[585] This implies that each citizen must possess principles and guidelines to which they can appeal in order to meet this standard: Rawls proposed that one way to identify such principles and guidelines is to show that they would have been agreed upon in what he refers to as the original position (Rawls, 1971, pp. 17–18, 1999, p. 140).[586] This is a hypothetical situation in which

---

all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason" (Rawls, 1993, p. 137).

[584] "When citizens share a reasonable political conception of justice they have a basis on which public discussion of fundamental political questions can proceed and be reasonably decided, not of course in all cases, but we hope in most cases of constitutional essentials and matters of basic justice" (Rawls, 1993, p. xxi). By "constitutional essentials," Rawls means the "fundamental principles that specify the general structure of government and the political process" and the "basic rights and liberties" of citizens (Rawls, 1993, p. 227). By "matters of basic justice," he means the chief structural features of the basic structure of society, namely economic justice, equality of opportunity, income, wealth, and minimum levels of welfare not already covered by the constitutional essentials (Rawls, 1993, pp. 228–229). In his last statement of justice as fairness, Rawls emphatically reaffirms this central idea: "We strive for publicly based justifications for questions regarding the constitutional essentials and basic questions of distributive justice but not in general for all the questions to be settled by the legislature within a constitutional framework. We should distinguish, then, between these two cases, the first attainable (we hope) and desirable, the second neither attainable nor desirable." (Rawls, 2001, p. 91). This is known as the "narrow view" of the scope of public reason, which Rawls contrast with a "wide view" (Rawls, 1997, p. 783).

[585] A bounded society is a self-sufficient and autonomous system of cooperation, whose members "enter only by birth and exit only by death" (Rawls, 1999, p. 136). As many have noticed, Rawls describes "bounded society" in terms very similar to a territorial state, even while explicitly denying that he meant "states" in the realist sense (O'Neill, 2001, p. 186). This "latent statism" has been criticized for being "strangely silent about the predicaments of outsiders, and about the justice of a world that is segmented into states, a world in which for each of us most others are emphatically not fellow-citizens." (O'Neill, 2000, p. 4).

[586] "A citizen engages in public reason, then, when he or she deliberates within a framework of what he or she sincerely regards as the most reasonable political conception of justice, a conception that expresses political values that others, as free and equal citizens might also reasonably be expected reasonably to endorse. Each of us must have principles and guidelines to which we appeal in such a way that this criterion is satisfied. I have proposed that one way to identify those political principles and guidelines is to show that they would be agreed to in what in Political Liberalism is called the original position. Others will think that different ways to identify these principles are more reasonable."(Rawls, 1999, p. 140). This formulation already exposes some of the tensions in Rawls's account, for example regarding the role that sincerity plays in it: how should sincerity be defined, and by what means verified? For

rational individuals come together to choose the fundamental principles that will govern a just society. What makes this situation unique is that the individuals are placed behind a veil of ignorance: they have no knowledge of their own personal characteristics, such as their class, race, gender, talents, or life goals (Rawls, 1971, p. 18). This ensures that the principles they select are fair and impartial, since no one can tailor the rules to benefit their own specific position in society.

To clarify how these principles selected in the original position would be applied in practice, Rawls introduces the "four-stage sequence" (Rawls, 1971, p. 195), which allows for the gradual introduction of information and social context while maintaining the fairness ensured by the original position. At each stage, the parties remain committed to the principles of justice and are prevented from acting in self-interest by the continuing effects of the veil of ignorance. The first stage is the constitutional convention, where parties are tasked with designing a just political constitution. They determine the basic rights and duties of citizens and establish the framework of government. Although they now consider general facts about society and political institutions, they still lack knowledge of their own position within it, ensuring impartiality. The second stage involves "legislation". Here, the parties take on the role of legislators, creating laws within the framework of the previously chosen constitution. They aim to realize the two principles of justice through legal structures and policies. While they now have access to more detailed social information—such as economic theory and general facts about human behavior—they still do not know which social class or group they belong to (Rawls, 1971, p. 199).[587]

In the third stage, Rawls introduces the application of rules to particular cases. The focus shifts to administrative and judicial officials who must implement and interpret laws in ways that remain faithful to the agreed-upon principles. By this point, the parties have a clearer picture of

Rawls, sincerity seems intended to secure stability through reasonable justification, yet it risks being an overly demanding and perhaps ineffective condition.

[587] The second principle of justice (the fair equality of opportunity and the difference principle) comes into play at this stage of legislation, while the first principle of justice (the principle of equal liberty) comes into play at the first stage of the constitutional convention. "I imagine then a division of labor between stages in which each deals with different questions of social justice. This division roughly corresponds to the two parts of the basic structure. The first principle of equal liberty is the primary standard for the constitutional convention. Its main requirements are that the fundamental liberties of the person and liberty of conscience and freedom of thought be protected and that the political process as a whole be a just procedure. Thus the constitution establishes a secure common status of equal citizen ship and realizes political justice. The second principle comes into play at the stage of the legislature. It dictates that social and economic policies be aimed at maximizing the long-term expectations of the least advantaged under conditions of fair equality of opportunity, subject to the equal liberties being maintained." (Rawls, 1971, p. 199).

how society functions and how the laws affect different groups, but their personal identity remains hidden. Finally, in the judicial stage, the principles and laws are interpreted by judges in actual legal cases. This stage allows for the refinement of justice as it is applied to complex, real-world situations. The emphasis is on ensuring consistency, fairness, and fidelity to the constitutional and legislative framework developed in the earlier stages (Rawls, 1971, p. 199)

Throughout these four stages, Rawls emphasizes a careful balance between the progressive introduction of information and the preservation of impartiality (Rawls, 1999, p. 200). As the sequence unfolds, the flow of information gradually increases, allowing the parties to make more context-sensitive decisions; however, the veil of ignorance is never fully lifted. At no point knowledge of their own personal characteristics, social status, or particular life goals become relevant, because these are the kinds of information that could lead to biased or self-serving decisions. This ongoing limitation is essential: it ensures that all choices made throughout the sequence continue to reflect a commitment to fairness and equal respect for all citizens. In this way, the four-stage sequence serves to operationalize the original position by linking abstract principles to the concrete processes of political and legal decision-making (Rawls, 1999, p. 200).

Rawls explicitly conceives this procedure as Kantian in character, and he interprets the principles of justice as analogous to categorical imperatives in Kant's sense, not because they are general and universal, but because they are the expression of the interests of free and equal moral persons (Rawls, 1971, pp. 251–253). Therefore for Rawls to act according to the principles of justice means to act according to categorical imperatives, since they apply independently of our particular ends; he points out that the veil of ignorance deprives the choosers of information about themselves that would "enable them to choose heteronomous principles," because parties in the original position lack any "particular desires" and have only a "preference for primary goods […] derived from only the most general assumptions about rationality and the conditions of human life" (Rawls, 1971, p. 253). He notes that his assumption of the "mutual disinterest" of the choosers' motivations can be seen as a procedural interpretation of the link between the Kantian conception of autonomy and the categorical imperative: the principles that govern the "kingdom of ends" are those that would be chosen in this position (Rawls, 1971, p. 256).[588] The description

[588] "The original position may be viewed, then, as a procedural interpretation of Kant's conception of autonomy and the categorical imperative. The principles regulative of the kingdom of ends are those that would be

of the original position makes it possible to explain the sense in which acting according to these principles expresses our nature as free and equal rational persons. In this way, notions that might otherwise appear purely transcendent and disconnected from human conduct acquire concrete relevance through the procedural conception of the original position

Many Kantian scholars, particularly those in the constructivist tradition, have emphasized the significant ways in which Rawls's original position departs from the Kantian framework (Habermas, 1995; O'Neill, 2000). While we will explore the critiques of Habermas and O'Neill in greater detail later in this and the next section, it is worth briefly highlighting some of their main concerns regarding the claim that Rawls's procedure is fully Kantian. Habermas believes that even if Rawls attempted to avoid philosophical commitments by relying on the "reasonable" instead of the "true", the procedure of the original position still incorporates morally substantive considerations (Habermas, 1998, p. 99).[589] This is not necessarily a problem—in fact, it is close in spirit to Kant, the problem is that Rawls seeks to avoid this outcome, but he ultimately fails to adequately model impartiality or to represent the "higher-order interests" of real citizens (Habermas, 1998, p. 52). For Habermas, a context-independent procedure of impartial judgment remains essential for articulating the moral point of view. Such a procedure need not to be neutral regarding normative reason, but only neutral regarding arbitrary values. In Rawls' proposal these elements are concealed and blurred together: although citizens can agree on a political conception (from the perspective of their several comprehensive doctrines), they do not agree for the same reasons. What is missing for Habermas is a common perspective on justice among citizens and a genuine moral consensus.

chosen in this position, and the description of this situation enables us to explain the sense in which acting from these principles expresses our nature as free and equal rational persons." (Rawls, 1971, p. 256).

[589] "The problematic relation between the reasonable and the true calls for an explanation that raises questions concerning Rawls's strategy of avoidance. The concept of practical reason cannot be drained of moral substance and morality cannot be relegated to the black box of comprehensive doctrines. I cannot see any plausible alternative to the straightforward Kantian strategy. There seems to be no way around the explanation of the moral point of view in terms of a procedure that claims to be context-independent. Such a procedure is by no means free of normative implications" (Habermas, 1998, p. 99).

Similarly, O'Neill accuses Rawls of departing significantly from Kant (O'Neill, 1989, pp. 207–208).[590] Whereas Kantian theories are fundamentally grounded in obligations, Rawls's approach emphasizes individual entitlements derived from a hypothetical social contract. In doing so, according to O'Neill, Rawls prioritizes what individuals are owed over what they owe to others, thereby weakening the relational and reciprocal dimensions of moral agency that are central to Kant's ethics. (O'Neill, 1989, p. 214).[591] This emphasis on entitlement is tied to Rawls's reliance on the hypothetical consensus under the idealizing conditions of the original position. However, the "clever packaging of the veil of ignorance" reflects a particular "ideal of the human subject" that is not purified of normative assumptions (O'Neill, 2000, p. 71),[592] and some of these assumptions are not properly Kantians (O'Neill, 2000, p. 45).[593] One of her central criticisms is that Rawls' account ignore the actual vulnerabilities of embodied human beings: the key-elements of plurality, connection, and finitude (O'Neill, 1996, p. 100) are unjustifiably absent from his idealized conception of the agent. O'Neill challenges this move by arguing that genuine moral principles must be justifiable not merely to idealized agents: they need to be accessible and potentially acceptable to those who are bound by their conclusions (O'Neill, 1989, p. 209).[594]

---

[590] "Without doubt important aspects of A Theory of Justice are Kantian. However, Rawls departs fundamentally from Kant by relying on a solely instrumental conception of rationality. Justice constructed on this basis must seemingly consist of principles that Kant would have rejected as heteronomous. This danger is to be averted by shrouding the original position in a veil of ignorance (TJ, 252). Those who do not know what they want cannot choose self-serving principles. However, instrumental reasoning is wholly at sea unless oriented by some goals. Hence Rawls lifts the veil enough to allow knowledge that certain "primary" — rights and liberties, powers and opportunities, income and wealth (TJ, 62) — are universally desired. This filtered access to knowledge of desires is carefully limited to ensure that parties in the original position are mutually disinterested." (O'Neill, 1989, pp. 207–208). In short, as O'Neill reiterates in several of her writings, justice as fairness involves a different kind of procedure, and has a different sphere of application, from Kant's categorical imperative (O'Neill, 2015a, pp. 74–75).

[591] "Rawls' focus on entitlement rather than on obligation is significant. Constructing a set of entitlements is not the same as constructing principles of obligation." (O'Neill, 1989, p. 214).

[592] "My own belief is that idealizations are common in the background assumptions of contemporary liberal, Kantian theories of justice, and that they are not always explicitly justified. I shall illustrate rather than document this by indicating some features of the work of John Rawls." (O'Neill, 2000, p. 71). O'Neill draws a crucial distinction between idealization and abstraction, accusing Rawls of relying on the former—a point we will explore in detail in the next section.

[593] "Reluctance to credit human agents with noumenal freedom lies behind Rawls's reluctance to go the whole way with Kant, and Sandel's polemic against something that he calls the 'deontological self' and attributes to Kant and to contemporary 'Kantians' (O'Neill, 2000, p. 45). Another significant divergence from Kant's notions of autonomy and reason lies in the fact that Rawls "departed fundamentally from Kant by relying on a solely instrumental conception of rationality" (O'Neill, 1988, p. 3).

[594] By detecting a direct link between abstract principles and their practical justification, O'Neill advocates for a conception of justice that maintains a delicate balance between universality and particularity: the universality of

That the original position is a distinct procedure from Kant's categorical imperative is something that even Rawls partially acknowledged in his later writings, such as *"Kantian Constructivism in Moral Theory"*. There, while he still argues that there is an analogy between justice as fairness and Kant's moral philosophy in certain fundamental respects, he places considerably less emphasis on the specific parallel between the categorical imperative and the original position (Rawls, 1980, p. 535); instead, he shifts the focus to how the two moral powers of the agent—a sense of justice and a conception of the good—are embedded in the design of the original position (Rawls, 1980, pp. 543–544). In doing so, Rawls is trying to underline the importance of replacing the search for moral truth—understood as the discovery of an independent order of moral facts and relations—with the search for reasonable bases for agreement. In this framework, moral universality is understood not in terms of correspondence to an independent moral reality, but in terms of a "suitably constructed social point of view" that all can endorse: outside the procedure for constructing principles of justice, he contends, there are no independent moral facts (Rawls, 1980, p. 519). The procedure is not a deduction from self-evident truths, but a practical task of constructing universal truths.

A crucial part of this practical procedure is the Reflective equilibrium, which is both a method and a state of coherence among our beliefs about justice, involving a back-and-forth adjustment between general assumption and specific considerations to achieve considered judgments (Rawls, 1980, p. 534).[595] If a theory aligns well with our considered judgments and

---

moral principles must not come at the expense of their intelligibility and relevance to the moral lives of real persons. This stance reflects her broader commitment to a form of constructivism that is less reliant on idealization and more attentive to the moral and epistemic limits of human agents. Rawls's framework, by contrast, risks constructing a model of justice that is too far removed from the lived experiences, vulnerabilities, and interdependencies that characterize human moral life. We are not isolated rational agents, but embodied, interdependent, and "morally fallible beings" (O'Neill, 1989, p. 209)

[595] Several scholars emphasize the centrality of reflective equilibrium throughout Rawls's procedure, viewing it as fundamental not only to the selection of the principles of justice within the original position but also to the broader justification of justice as fairness as a comprehensive conception (Finlayson, 2019, p. 24). However, in this text, Rawls distinguishes three distinct points of view within his theory: 1)The perspective of the parties in the original position, who represent rational agents choosing principles under specified constraints; 2) The standpoint of citizens in a well-ordered society, who embody the moral ideal of full autonomy; 3) The viewpoint of us—the theorists and evaluators, who assess justice as fairness from our present reflective standpoint. Rawls appears, thus, to suggest that the method of wide reflective equilibrium operates primarily—if not exclusively—at this third level, where the overall coherence of the theory is tested against our considered moral judgments across various levels of generality. "The third point of view-that of you and me-is that from which justice as fairness, and indeed any other doctrine, is to be assessed. Here the test is that of general and wide reflective equilibrium, that is, how well the view as a whole meshes with and articulates our more firm considered convictions, at all levels of generality, after due examination, once all adjustments

enables us to express what we wish to affirm, then it provides a way to replace familiar formulations with better ones. Reflective equilibrium, as such, is central to both constructivism and rational intuitionism; the key difference lies, however, in how they interpret unacceptable outcomes: for the intuitionist, such results indicate an error in the independent order of values; for the constructivist, the error must lie in how the procedure shapes the principles of practical reason (Rawls, 1980, p. 570).[596] It is here that the priority of the right over the good, and of the reasonable over the rational—in the senses defined by Rawls—is affirmed as a defining feature of Kantian constructivism (Rawls, 1980, p. 532).[597]

Subsequently, in *Justice as Fairness: Political Not Metaphysical* (1985), Rawls increasingly emphasized a point that would be further developed in *Political Liberalism* (1993): that a theory of justice with liberal content is best understood as a political conception. A political conception of justice is justified by appeal to political values and should not be presented as part of anymore "comprehensive" moral, religious, or philosophical doctrine (Rawls, 1985, pp. 225–226, 1993, p. 12).[598] Rawls ceases to refer to notions such as moral truths or moral universality,

---

and revisions that seem compelling have been made. A doctrine that meets this criterion is the doctrine that, so far as we can now ascertain, is the most reason able for us." (Rawls, 1980, p. 534).

[596] "My account of Kantian constructivism in moral theory (as illustrated by justice as fairness) is now concluded. I should stress, however, that for all I have said it is still open to the rational intuitionist to reply that I have not shown that rational intuitionism is false or that it is not a possible basis for the necessary agreement in our judgments of justice. It has been my intention to describe constructivism by contrast and not to defend it, much less to argue that rational intuitionism is mistaken. In any case, Kantian constructivism, as I would state it, aims to establish only that the rational intuitionist notion of objectivity is unnecessary for objectivity. Of course, it is always possible to say, if we ever do reach general and wide reflective equilibrium, that now at last we intuit the moral truths fixed by a given moral order; but the constructivist will say instead that our conception of justice, by all the criteria we can think of to apply, is now the most reasonable for us." (Rawls, 1980, p. 570).

[597] The Reasonable (public reasoning about fair terms of cooperation) frames and limits the Rational (individual reasoning about how best to achieve one's own ends or interests) That is, moral principles (the Reasonable) set the boundaries within which ethical values (the Rational) can take place. "Finally, the way in which the Reasonable frames the Rational in the original position represents a feature of the unity of practical reason. In Kant's terms, empirical practical reason is represented by the rational deliberations of the parties; pure practical reason is represented by the constraints within which these deliberations take place. The unity of practical reason is expressed by defining the Reasonable to frame the Rational and to subordinate it absolutely; that is, the principles of justice that are agreed to are lexically prior in their application in a well-ordered society to claims of the good. This means, among other things, that the principles of justice and the rights and liberties they define cannot, in such a society, be overridden by considerations of efficiency and a greater net balance of social well-being. This illustrates one feature of the unity of reason: the Reasonable and the Rational are unified within one scheme of practical reasoning which establishes the strict priority of the Rea sonable with respect to the Rational. This priority of the right over the good is characteristic of Kantian constructivism." (Rawls, 1980, p. 532).

[598] The theory of justice is now explicitly formulated to apply only to the basic structure of society—that is, to the main political, social, and economic institutions. It is not intended as a general moral doctrine governing all

favoring instead the concept of the reasonable; however, unlike in earlier formulations, the term now takes on a technical meaning (Finlayson, 2019, p. 113),[599] and refers to a conception that can be endorsed and sustained by an overlapping consensus among citizens who hold diverse comprehensive doctrines, without needing to be derived from or supported by any single doctrine considered "true." At this point in his theory, the notion of reasonable pluralism—not merely the "fact of pluralism"—begins to play a central role. A modern democratic society, he argues, is characterized not simply by the coexistence of diverse comprehensive religious, philosophical, and moral doctrines, but by a pluralism of incompatible yet reasonable comprehensive doctrines (Rawls, 1993, pp. xvi, xxvi).[600] At the same time, the concept of overlapping consensus becomes crucial too (Rawls, 1993, Chapter I, IV);[601] it consists in the endorsement, by different reasonable comprehensive doctrines, of the same political conception of justice, even if each doctrine supports this political conception based on its own internal grounds and reasons (Rawls, 1993, p. xlvii).[602]

---

individual actions, nor as a comprehensive account of global justice. "In this respect a political conception of justice differs from many moral doctrines, for these are widely regarded as general and comprehensive views. Utilitarianism is a familiar example: the principle of utility, however understood, is usually said to hold for all kinds of subjects ranging from the conduct of individuals and personal relations to the organization of society as a whole as well as to the law of peoples. By contrast, a political conception tries to elaborate a reasonable conception for the basic structure alone and involves, so far as possible, no wider commitment to any other doctrine." (Rawls, 1993, p. 12). This restriction of scope is essential, according to Rawls, to its political nature and to its capacity to gain legitimacy in a context of reasonable pluralism, as we will see in the next pages.

[599] "In A Theory of Justice, Rawls uses the term "rational" widely, sometimes as a semi-technical term, but he uses the word "reasonable" in its everyday sense. In Political Liberalism, the term "reasonable" becomes a term of art that bears much of the weight of his argument. In the later work, "reasonableness" is both a virtue that pertains to persons and a quality that a principle, or a body of principles, such as a comprehensive doctrine, may have or may lack." (Finlayson, 2019, p. 113). Similarly, the contrast between the metaphysical and the political is articulated through the use of these concepts as terms of art. A conception of justice is not metaphysical insofar as it does not rely on a comprehensive philosophical doctrine that reasonable citizens might reject. Were it to do so, it could not serve as the basis for a consensus among free and equal citizens.

[600] The fact of pluralism is just the fact that the sources of value are many and often conflicting and incommensurable. The fact of reasonable pluralism is the fact that this is also true of the subset of reasonable doctrines espoused by reasonable citizens, and that this is a "natural outcome". "For this reason, political liberalism assumes the fact of reasonable pluralism as a pluralism of comprehensive doctrines, including both religious and nonreligious doctrines. This pluralism is not seen as disaster but rather as the natural outcome of the activities of human reason under enduring free institutions." (Rawls, 1993, p. xxvi).

[601] Although the idea is already present in *A Theory of Justice* (Rawls, 1971, pp. 388, 476), it appears only twice—and mostly toward the end—whereas in *Political Liberalism* the term appears approximately 200 times and plays a central role throughout the entire exposition (Rawls, 1993, Chapter I, IV).

[602] The relation with overlapping consensus and reasonable pluralism is therefore subtle but crucial: reasonable pluralism is a sociological fact, whereas overlapping consensus is a theoretical construct and a practical objective aimed at addressing that fact in a just and stable way. "I shall not try to describe here the idea of an overlapping consensus beyond what I have occasionally said. I note instead two points related to it. One is that it is

To summarize and clarify our discussion thus far in light of our research objectives, I am trying to elucidate how Rawls gradually shifted from the concept of universality to that of objectivity, which he regarded as less metaphysically burdensome.[603] In A Theory of Justice, his approach is analytical and conceptual: competing conceptions of justice are evaluated according to how well they reflect the core concept of justice—namely, the rejection of arbitrary distinctions among persons (Rawls, 1971, p. 6).[604] However, ambiguity remained as to whether "justice as fairness" was intended as a comprehensive moral doctrine or as a political conception of justice. This ambiguity is explicitly resolved in Political Liberalism, where Rawls presents "justice as fairness" clearly as a political—not metaphysical—conception of justice. To support this shift, he foregrounds two ideas that had been only marginally present in his earlier work: (1) that a certain level of agreement on principles of justice is necessary for the stability of any human community; and (2) that the justification of those principles must rest on an *overlapping consensus* among reasonable citizens.

---

the fact of reasonable pluralism that leads-at least me-to the idea of a political conception of justice and so to the idea of political liberalism. For rather than confronting religious and nonliberal doctrines with a comprehensive liberal philosophical doctrine, the thought is to formulate a liberal political conception that those nonliberal doctrines might be able to endorse. To find this political conception we do not look at known comprehensive doctrines with the aim of striking a balance or average between them, nor do we attempt to strike a compromise with a sufficient number of those doc trines actually existing in society by tailoring the political conception to fit them. Doing that appeals to the wrong idea of consensus and makes the political conception political in the wrong way (39ff.). Rather, we formulate a freestanding political conception having its own intrinsic (moral) political ideal expressed by the criterion of reciprocity." (Rawls, 1993, p. xlvii). The criterion of reciprocity in Rawls requires that the proposed terms of social cooperation be such that every citizen could reasonably accept them. Forst (2012, p. 88) further emphasizes the Habermasian dimension of reciprocity and mutual recognition. See also footnote n.694 and n.730

[603] This shift also helps explain why *The Cambridge Rawls Lexicon* includes an entry on "Objectivity" (Mandle & Reidy, 2014, pp. 571–573), but none on "universality," "universalism," or "universalizability."

[604] Rawls starts *A Theory of Justice* precisely by distinguishing between the concept of justice (an abstract idea broadly shared) and its conceptions (the concrete interpretations of the concept). Although there are multiple conceptions of justice—each specifying particular distributions of rights, duties, and social benefits—they converge on the fundamental idea that no arbitrary distinctions should be made between persons: this is the core of the concept of justice that all the conceptions must have in common (Rawls, 1971, p. 6). This is the sense in which Rawls's initial approach is conceptual: while there is only one concept of justice, there are multiple competing conceptions, which can be evaluated according to how well they represent that shared concept. Despite Rawls's later characterization, it is clear that his aim in *A Theory of Justice* was to present "justice as fairness" as the moral conception that best captures our intuitive sense of justice, especially when compared with the dominant utilitarian paradigm: society is just not when it maximizes overall happiness, but when its basic structure is arranged according to principles that free and equal persons would agree to under fair conditions (the original position).

Rawls's approach thus becomes more social and pragmatic.[605] He recognizes that enduring disagreement on major philosophical and moral questions is inevitable, given that reasonable pluralism is "the natural outcome of the activities of human reason under enduring free institutions" (Rawls, 1993, p. xxvi). Rawls therefore no longer seeks to present his "justice as fairness" as the best representation of the concept of justice, nor to justify it through a comprehensive moral doctrine (such as those of Kant or Mill), as he believes this would constitute a metaphysical approach. For this reason, in Lecture III of *Political Liberalism*, he explicitly distinguishes between rational intuitionism, Kantian moral constructivism, and political constructivism (exemplified by his own justice as fairness), distinguishing their respective conceptions of objectivity. Here, he undertakes a comparative assessment of these moral frameworks by identifying six criteria that any adequate conception of objectivity must satisfy (Rawls, 1993, pp. 110–130).[606]

Rawls concludes that, although all three frameworks can meet the essential conditions for objectivity, they do so in fundamentally different ways. For example, with regard to the second criterion, in rational intuitionism correctness is defined in terms of truth about an independent moral order. By contrast, in Kantian and political constructivism, correctness is grounded in reasonableness—that is, in conformity with procedures of practical reason. What distinguishes Kantian moral constructivism from Rawlsian political constructivism, however, is that Rawls now

[605] Here I am not referring to a pragmatist conception of objectivity in the vein of Rorty, who replaces tout-court the pursuit of objectivity with the pursuit of mutual understanding, and therefore defines objectivity as "solidarity" (Rorty, 1989, 1990). Rawls is more directly influenced by Dewey's conception of "operative objectivity", which is validated through the consensus of the investigative community. This is particularly interesting in light of the fact that Dewey still retains a notion of universality—though not one that is abstract or beyond time and space, but rather one that "works" for anyone under comparable conditions (Dewey, 1925).

[606] The six criteria are the following. 1) Public Framework for Judgment: A conception of objectivity must offer a shared framework of reasoning within which individuals can make judgments based on reasons and evidence, after due reflection. This is a precondition for meaningful moral, political, or scientific inquiry. 2) Standard of Correctness: Objectivity requires a normative standard by which judgments can be evaluated as correct. 3) Normative Force of Reasons: Objective frameworks assign reasons to agents, not merely as motivations but as normative imperatives that may override personal inclinations. These reasons function within a system of principles that agents are expected to endorse and act upon. 4) Distinction between Objective and Subjective Standpoints: Objectivity requires distinguishing the agent's or group's point of view from a general standpoint accessible to all reasonable persons. The belief that something is just does not make it just; public justification must transcend individual or collective bias. 5) Possibility of Agreement among Reasonable Agents: A viable conception of objectivity must explain how reasonable agents, informed by the same facts and principles, can converge in judgment through reflection and deliberation. 6) Explanation of Disagreement: The theory must also provide a coherent account of the conditions under which disagreement among reasonable agents may arise and how such disagreement can be rationally understood. These six are all procedural features of practical reasons (Rawls, 1993, p. 114; Habermas, 1998, p. 96).

views the former as a comprehensive moral doctrine, which renders it incompatible with the aims of political liberalism. While he praises Kantian constructivism for securing objectivity not by appealing to metaphysical truths but through intersubjective agreement among rational persons, via a procedural conception that satisfies all essential requirements for a reasonable and rational objectivity (Rawls, 1993, p. 114),[607] he also argues that once extended beyond the political domain, Kantian constructivism becomes too metaphysical and as such it may not be shared in a pluralistic society; instead political liberalism, as Rawls envisions it, restricts itself to the political domain, deliberately avoiding contentious moral foundations to focus on what reasonable citizens can affirm in common (Rawls, 1993, p. 116).[608]

In this framework, coherence—not foundationalism—is the goal, achieved by integrating diverse moral judgments (e.g., against tyranny, slavery and exploitation) into a unified normative vision (Rawls, 1993, pp. 123–124).[609] This unifying feature however is achieved no longer through universality, but on its capacity to serve as the "focus of an overlapping consensus". This does not mean that there are no moral ground upon which the consensus is built, but only that the political conception of justice can be supported by various reasonable comprehensive doctrines, each for its own reasons, without any one doctrine needing to prevail (Rawls, 1993, pp. 147–148).[610] This

[607] "As for Kant's moral constructivism (§2), a correct moral judgment is one that meets all the relevant criteria of reasonableness and rationality incorporated into the categorical imperative procedure for testing maxims. A judgment properly supported by principles and precepts that pass that test are acknowledged as correct by any fully reasonable and rational (and informed) person. This is what Kant means when he says that these judgments are universally communicable: as reasonable and rational we recognize, apply, and can explain to others, the same procedure for validating them. All the essential elements of objectivity are provided for." (Rawls, 1993, p. 114).

[608] "I said above that the essentials of objectivity (including the sixth in §7.2) are necessary for a shared public basis of justification. To this add that they are also sufficient. With this, political liberalism has an account of objectivity that suffices for the purposes of a political conception of justice. As we have said, it need not go beyond its conception of a reasonable judgment and may leave the concept of a true moral judgment to comprehensive doctrines." (Rawls, 1993, p. 116)

[609] "Some may ask: why look for something to ground the fact that slavery is unjust? What is wrong with the trivial answer: slavery is unjust because slavery is unjust? Can't we stop with that? Why talk about possibilities of construction? Political constructivism does not look for something for the reasonableness of the statement that slavery is unjust to consist in, as if the reasonableness of it needed some kind of grounding. We may accept provisionally, though with confidence, certain considered judgments as fixed points, as what we take as basic facts, such as slavery is unjust. But we have a fully philosophical political conception only when such facts are coherently connected together by concepts and principles acceptable to us on due reflection. These basic facts do not lie around here and there like so many isolated bits. For there is: tyranny is unjust, exploitation is unjust, religious persecution is unjust, and on and on. We try to organize these indefinitely many facts into a conception of justice by the principles that issue from a reasonable procedure of construction" (Rawls, 1993, pp. 123–124).

[610] This consensus is more stable than a mere modus vivendi, precisely because it is based on a moral commitment to the political conception itself. "That an overlapping consensus is quite different from a modus vivendi

leads to a central feature of political liberalism: the recognition of a diversity of liberal political conceptions of justice and "therefore many forms of public reason specified by a family of reasonable political." (Rawls, 1997, p. 774).[611] Rawls conception of "justice as fairness" is just one example of such a liberal conception, and its specific content is not the only possible one: Political liberalism is broad enough to include different perspectives, like Habermas's discursive conception (Rawls, 1999, p. 142).[612]

5.1.2 Habermas's Discourse Theory and Principle (U)

Despite their common Kantian heritage and intellectual affinities, Habermas — in contrast to Rawls — preserves a central role for universality within his theory of communicative action.

---

is clear from our model case. In that example, note two aspects: first, the object of consensus, the political conception of justice, is itself a moral conception. And second, it is affirmed on moral grounds, that is, it includes conceptions of society and of citizens as persons, as well as principles of justice, and an account of the political virtues through which those principles are embodied in human character and expressed in public life. An overlapping consensus, therefore, is not merely a consensus on accepting certain authorities, or on complying with certain institutional arrangements, founded on a convergence of self- or group interests. All those who affirm the political conception start from within their own comprehensive view and draw on the religious, philosophical, and moral grounds it provides. The fact that people affirm the same political conception on those grounds does not make their affirming it any less religious philosophical, or moral, as the case may be, since the grounds sincerely held determine the nature of their affirmation." (Rawls, 1993, pp. 147–148).

[611] This commitment to public reason is extended in *The Law of Peoples*, where Rawls applies the framework of political liberalism beyond the domestic context. He distinguishes between the public reason of equal citizens within a liberal society and that of equal *peoples* within the international sphere. Although these two forms of public reason differ in content, their normative function is analogous: each is meant to guide deliberation among free and equal agents—citizens domestically, peoples internationally—on matters of justice and political legitimacy (Rawls, 1999, p. 55). As he explicitly puts it: "it is important to understand that the Law of Peoples is developed within political liberalism. This beginning point means that the Law of Peoples is an extension of a liberal conception of justice for a domestic regime to a Society of Peoples. Developing the Law of Peoples within a liberal conception of justice, we work out the ideals and principles of the foreign policy of a reasonably just liberal people. I distinguish between the public reason of liberal peoples and the public reason of the Society of Peoples. The first is the public reason of equal citizens of domestic society debating the constitutional essentials and matters of basic justice concerning their own government; the second is the public reason of free and equal liberal peoples debating their mutual relations as peoples. The Law of Peoples with its political concepts and principles, ideals and criteria, is the content of this latter public reason. Although these two public reasons do not have the same content, the role of public reason among free and equal peoples is analogous to its role in a constitutional democratic regime among free and equal citizens." (Rawls, 1999, p. 55). Scholars (Pogge, 1989; Beitz, 2000) criticizes Rawls for failing to apply his own moral logic consistently at the global level, arguing that this failure stems precisely from prioritizing *peoples* over individual *persons*.

[612] "Political liberalism, then, does not try to fix public reason once and for all in the form of one favored political conception of justice. That would not be a sensible approach. For instance, political liberalism also admits Habermas's discourse conception of legitimacy (sometimes said to be radically democratic rather than liberal), as well as Catholic views of the common good and solidarity when they are expressed in terms of political values." (Rawls, 1999, p. 142).

To fully grasp the implications of this theoretical move, it is helpful to revisit their respective conceptions of public reason, this time through the lens of Habermas's nuanced praise and critique (Habermas, 1995, p. 109).[613] Habermas faults Rawls for presenting the "original position" as a monological procedure that inadequately addresses moral autonomy underlying and is therefore unable to deliver a truly dialogical procedure that express the public capacity of reason, not of a philosophical expert (Habermas, 1990, p. 74,122).[614] The stylized choosers in the original positions are too abstract to represent real moral agents, unlike participants in discourse ethics, who are engaged in actual dialogical processes (Habermas, 1995, p. 117); moreover, the idea that the veil of ignorance ensures impartiality is misplaced, since impartiality must be continuously reaffirmed through public discourse, which the original position does not require (Habermas, 1995, pp. 116–19). In this way, Rawls limited the domain of "public reason" only to "constitutional essentials and matters of basic justice," too restrictive and artificial since morality "crosses the boundaries between private and public spheres" (Habermas, 1996, pp. 109–110).[615]

---

[613] I will primarily focus on *Between Facts and Norms* (1996), due to both the centrality of its subject matter and its substantial length. However, with respect to Habermas's engagement with Rawls and his critique of the latter's conception of public reason, two additional texts deserve specific consideration: *Moral Consciousness and Communicative Action* (1990) and the essay *Reconciliation through the Public Use of Reason: Remarks on John Rawls's Political Liberalism* (1995). It is also worth emphasizing here how Habermas frames his critique in this direct text: not as an external or oppositional stance, but as a difference of opinion within a shared intellectual project—a dispute within a family. As he puts it: "Because I admire this project, share its intentions, and regard its essential results as correct, the dissent I express here will remain within the bounds of a familial dispute." (Habermas, 1995, p. 110).

[614] As Finlayson (2019, pp. 52–55) emphasizes, two important aspects of Habermas's critique, though seemingly in tension, are in fact closely interrelated here. First, Habermas maintains that Rawls adopts a substantive approach by prescribing specific principles of justice from the perspective of a philosophical expert. In contrast, discourse ethics is procedural, insofar as it allows moral norms to emerge from the deliberation of participants themselves. Second, Habermas insists on a sharper methodological distinction between the role of the theorist and that of the participant. Within discourse ethics, the philosopher engages not as an authoritative legislator, but as one interlocutor among others in the shared practice of reason-giving.

[615] "Discourse theory conceives of morality as an authority that crosses the boundaries between private and public spheres; these boundaries vary throughout history anyhow, depending on social structure. If we construe the universalistic claim of the moral principle intersubjectively, then we must relocate ideal role taking, which according to Kant each individual undertakes privately, to a public practice implemented by all person in common." (Habermas, 1996, pp. 109–110). Still, Habermas has also been criticized for holding a narrow, even monolithic, conception of public reason. Notably, Nancy Fraser challenges his notion of the public sphere, arguing that it overlooks the reality of multiple, overlapping publics. According to Fraser, the public sphere is always composed of subaltern counterpublics, "parallel discursive arenas where members of subordinated social groups invent and circulate counterdiscourses, which in turn permit them to formulate oppositional interpretations of their identities, interests, and needs" (Fraser, 1990, p. 67). Therefore, while Habermas's ideal aimed to promote rational and open public deliberation, it was for Fraser nonetheless grounded in significant exclusions—particularly along the lines of gender, class, and private interest. The bourgeois public sphere, as Fraser notes, was defined in contrast to more "feminine" or "private" forms of culture, embedding masculinist norms that effectively excluded women from political

To respond to these limitations, Habermas introduces a more systematic distinction between "morality" and "ethics" (or moral discourse and ethical discourse), and clarifies also their relationship to political and legal discourse (Habermas, 1996, p. 109).[616] On this basis, he develops the idea of the "co-originality" of the public and private use of reason (which is reflected also in the political co-originality of individual rights and democratic participation) by offering a genealogical account of the emergence of secular morality, grounded in processes of rationalization that have progressively replaced the normative authority of religious and metaphysical worldviews. As the "lifeworld" becomes increasingly rationalized (Habermas, 1996, p. 302),[617] "the normative substance of an ethos embedded in religious and metaphysical traditions" is subjected to post-traditional justification, shifting the burden of social integration and norms legitimacy to communicative rationality and discourse ethics (Habermas, 1996, p. 99).[618]

---

participation. For Fraser, then, a genuinely democratic public sphere cannot merely suspend social inequalities, as the bourgeois model aspired to do; it must actively work to dismantle them.

[616] A crucial source of misunderstanding in the debate between Habermas and Rawls lies in their divergent conceptions of the "political" and of the relationship between morality and politics. Rawls adopts a more restricted view, treating political values as a relatively fixed and identifiable subset of general moral values. In contrast, Habermas envisions morality as permeating the political domain, in a way that "flows into and around the political system," thereby emphasizing a more dynamic and procedural interrelation between the two spheres (Finlayson, 2019, p. 138). Habermas's misunderstanding is based on the assumption that Rawls's political conception relies solely on contingent overlaps among citizens' private comprehensive doctrines, overlooking that Rawls presents it as a detachable module that fits within any reasonable doctrine and that the overlapping political values are moral in nature. However, even after correcting this misunderstanding, a substantive objection remains: Rawls draws an overly narrow boundary around the "political"—limiting public reason to constitutional essentials and basic justice—which arguably detaches his concept of the political from real-world political practices and sociological reality. This interpretation is backed up by the fact that some of Rawls's staunchest defenders, and some of his most sympathetic critics, attribute this narrow conception of public reason to him (Freeman, 2000, p. 405).

[617] Habermas uses the phenomenological notion of lifeworld, coined by Husserl, to refers to the background context of pre-reflexive shared understandings, meanings, and cultural traditions that individuals rely on in everyday social interactions. "The political system depends on the performance of other systems, such as the fiscal performances of the economic system, in more than just a trivial manner. What is more, deliberative politics is internally connected with contexts of a rationalized lifeworld that meets it halfway. This is true both for the politics governed by the formal procedures of an institutionalized opinion- and will-formation and for the politics that occurs only informally in the networks of the public sphere. It is precisely the deliberatively filtered political communications that depend on lifeworld resources-on a liberal political culture and an enlightened political socialization, above all on the initiatives of opinion building associations. To a large extent, these resources form and regenerate spontaneously, and in any case they are not readily accessible to direct interventions of the political apparatus." (Habermas, 1996, p. 302).

[618] Habermas identifies two historically dominant and competing traditions of political justification that attempt to ground the legitimacy of law under conditions of modernity: 1) the liberal tradition—exemplified by Kant and natural law theorists— which grounds legitimacy in moral autonomy and human rights, which are taken to precede and constrain democratic procedures (the universality of moral principles is independent of the political autonomy instantiated through the social contract); 2) the republican tradition—articulated most clearly by Rousseau—which bases political legitimacy on the sovereign will of the people, expressed through the general will and realized via a

Habermas, in short, starts from a specific practical issue of modern societies (Finlayson, 2019, p. 28), in which the social order can no longer rely on preexisting traditions or metaphysical beliefs; instead, it must be continually negotiated and reconstructed and for Habermas the medium through which this task can be accomplished is communication and its reflective counterpart, discourse (Habermas, 1996, pp. 17–18).[619] It is important here to take a step back and distinguish between Habermas's notions of communication and discourse, as the two are often mistakenly treated as synonymous. Communication represents the primary mode through which individuals interact and strive to achieve mutual understanding, which is "the inherent telos of human speech." (Habermas, 1985, p. 287).[620] In this everyday practice, the three validity claims of truth, rightness, and veracity are generally accepted implicitly and without being thematized;[621] the success of

social contract understood as a transformative act of collective self-constitution (the culturally homogenous citizenry is endowed with civic virtues and tends to subordinate individual autonomy to collective ethical aims). Habermas argues that both traditions, while historically exhaustive, are normatively incomplete: liberalism fails to fully account for public autonomy and the democratic legitimacy of law, while republicanism risks the suppression of private autonomy in favor of collective self-rule. He therefore calls for a reconciliation of these tensions through a conception of democracy that affirms the co-originality of private and public autonomy, wherein individual rights and democratic participation are mutually constitutive and interdependent. (Habermas, 1996, p. 101). While this reconstruction may be criticized for its oversimplification—particularly with regard to Kant, as should be evident from the analysis in section 4.1.3 (Ripstein, 2009)—it nevertheless offers valuable insight into the genealogical development of Habermas's conception of the *co-originality* of private and public autonomy, clarifying the sense in which these two dimensions possess equal normative weight and priority.

[619] "Every social interaction that comes about without the exercise of manifest violence can be understood as a solution to the problem of how the action plans of several actors can be coordinated with each other in such a way that one party's actions "link up" with those of others. An ongoing connection of this sort reduces the possibilities of clashes among the doubly contingent decisions of participants to the point where intentions and actions can form more or less conflict-free networks, thus allowing behavior patterns and social order in general to emerge. As long as language is used only as a medium for transmitting information, action coordination proceeds through the mutual influence that actors exert on each other in a purposive-rational manner. On the other hand, as soon as the illocutionary forces of speech acts take on an action-coordinating role, language itself supplies the primary source of social integration. Only in this case should one speak of "communicative action."" (Habermas, 1996, pp. 17–18).

[620] This may appear to conflate the concept with that of the *lifeworld* as explained in the previous note; however, Habermas in fact maintains a clear distinction between the two. "We do not need to go into the formal-pragmatic analysis of the lifeworld here, nor into the place that communicative action occupies in a broader theoretical architectonic (i .e., between rational discourse and lifeworld). The lifeworld forms both the horizon for speech situations and the source of interpretations, while it in turn reproduces itself only through ongoing communicative actions. In the present context, what interests me about background knowledge is its peculiar pre-predicative and precategorial character" (Habermas, 1996, p. 22). The *lifeworld* is essentially the implicit (back)ground upon which everyday communication takes place.

[621] According to Habermas, every communicative speech act implicitly raises one of these three distinct validity claims, each corresponding to a different dimension of reality or what he calls a different "world". The first is the claim to Truth (*Wahrheit*) which pertains to the *objective world* and concerns descriptive assertions or propositions regarding states of affairs (Habermas, 1996, p. 13). When a speaker makes an assertion such as "the sky is blue" they implicitly claim that the content of the statement is true. This claim to truth is characterized by binary validity meaning a proposition is either true or false. Its justification is pursued within theoretical discourse where interlocutors aim to

communication depends precisely on the intersubjective recognition of these claims (Habermas, 1996, p. 4).[622] Discourse, by contrast, constitutes the reflective counterpart of communication. It does not represent the default form of interaction but rather emerges when ordinary communication breaks down. Specifically, whenever a validity claim is contested or no longer tacitly accepted by one of the participants, the interaction shifts from the domain of action to that of discourse or argumentation (Habermas, 1996, p. 15). Discourse is thus a reflective form of communication that operates on a higher level than action oriented toward mutual understanding: its function is to serve as a reflective medium in which problematic or disrupted consensus can be renewed or replaced when the shared understandings that normally coordinate social interactions are compromised.

In sum, communication is the spontaneous everyday interaction aimed at reaching understanding, whereas discourse is the reflective and structured response to disagreement or a breakdown of consensus within communication, through which an attempt is made to restore understanding on rational and shared grounds. Furthermore, Habermas structures this disciplined

---

achieve rationally motivated agreement about the correspondence between the assertion and the empirical facts of the world. The second is the claim to Rightness *(Richtigkeit)* which relates to the *social or intersubjective* world and concerns normative statements or norms (Habermas, 1998, p. 39). When one states for example "one must not steal" they are implicitly asserting the correctness or justifiability of the underlying norm. This claim too is governed by binary validity as norms are either valid or invalid. Justification takes place within practical or moral discourse where the goal is to reach consensus regarding the generalizability of the norm and its acceptability by all affected. Valid moral norms are those that incorporate generalizable interests and can be affirmed as equally good for all. The third is the claim to Veracity *(Wahrhaftigkeit)* which concerns the *subjective* world of the speaker and involves self-expressions including emotions, intentions, attitudes and values. When a speaker utters statements such as "I am happy" or "my intention is to do X" they implicitly claim to be sincere or authentic in their expression. Unlike the previous two claims, truthfulness is characterized by a gradational rather than binary validity as self-expressions may be more or less authentic or sincere. In this domain practical reflection assumes the form of hermeneutic self-clarification: veracity, understood as sincerity, is a fundamental precondition for genuine argumentation and meaningful interpersonal understanding. Finally, it is useful to underline, for our specific aim, how Habermas asserts that the three validity claims are "universal features of discourse" (Habermas, 1985, p. 75). Truth claims and rightness claims, specifically are universal in the sense that, if an assertion is true, it is true for everyone, and if a norm is right, it is right for everyone. There are no exceptions.

[622] "This rationality is inscribed in the linguistic telos of mutual understanding and forms an ensemble of conditions that both enable and limit. Whoever makes use of a natural language in order to come to an understanding with an addressee about something in the world is required to take a performative attitude and commit herself to certain presuppositions. In seeking to reach an understanding, natural-language users must assume, among other things, that the participants pursue their illocutionary goals without reservations, that they tie their agreement to the intersubjective recognition of criticizable validity claims, and that they are ready to take on the obligations resulting from consensus and relevant for further interaction. These aspects of validity that undergird speech are also imparted to the forms of life reproduced through communicative action. Communicative rationality is expressed in a decentered complex of pervasive, transcendentally enabling structural conditions, but it is not a subjective capacity that would tell actors what they ought to do. Unlike the classical form of practical reason, communicative reason is not an immediate source of prescriptions. It has a normative content only insofar as the communicatively acting individuals must commit themselves to pragmatic presuppositions of a counterfactual sort." (Habermas, 1996, p. 4).

practice of discourse through three distinct levels of rules. The first level includes basic logical and semantic principles, such as non-contradiction, universalizability, and consistency of meaning (Habermas, 1990, p. 86).[623] The second level consists of procedural norms like sincerity and accountability, requiring participants to assert only what they believe and to justify their claims when asked; these are preconditions for all genuine argumentation, that is, argumentation conceived as a "search for truth" (Habermas, 1990, pp. 87–88). The third level involves conditions that safeguard discourse from coercion, repression, and inequality, ensuring that only the "unforced force of the better argument" guides the process toward rational consensus (Habermas, 1998, p. 37);[624] such rules are designed to ensure that any consensus reached within discourse is rationally motivated (Habermas, 1990, p. 89, 1996, p. 103).[625]

From this brief outline, we can see how, for Habermas, universality is a first-order procedural criterion governing the practice of rational discourse, essential for the justification and validation of the rightness of norms. Habermas formalizes and calls this criterion principle of Universalization (U)."(U) A norm is valid if and only if the foreseeable consequences and side effects of its general observance for the interests and value-orientations of each individual could be freely accepted jointly by all concerned." (Habermas, 1998, p. 42). A key notion in principle (U) is that of a universalizable or generalizable interest. This means that a norm is justified only if it satisfies the interests of all those affected and is adopted by everyone because they recognize its

[623] "(1.1) No speaker may contradict himself. (1.2) Every speaker who applies predicate F to object A must be prepared to apply F to all other objects resembling A in all relevant respects. (1.3) Different speakers may not use the same expression with different meanings" (Habermas, 1990, p. 86).

[624] "[E]very justification muscat least proceed from a pre-understood context or background understanding. This failure of foundationalism recommends a pragmatic conception of justification as a public practice in which criticizable validity claims can be defended with good reasons. Of course, the criteria of rationality that determine which reasons count as good reasons can themselves be made a matter for discussion. Hence procedural characteristics of the process of argumentation itself must ultimately bear the burden of explaining why results achieved in a procedurally correct manner enjoy the presumption of validity. For example, the I communicative structure of rational discourse can ensure that all relevant contributions are heard and that the unforced force of the better argument alone determines the "yes" or "no" responses of the participants." (Habermas, 1998, p. 37).

[625] Habermas draws on the work of the German legal theorist Robert Alexy to illustrate examples of these third types of rules to which participants in discourse are always already committed. He provides three notable examples: 1) Every subject capable of speech and action is entitled to participate in discourse; 2) All participants are permitted to: a. Question any assertion whatsoever; b. Introduce any assertion into the discourse; c. Express their attitudes, desires, and needs. 3) No speaker may be prevented—whether by internal or external coercion—from exercising the rights granted under (1) and (2) (Habermas, 1990, p. 89). These are not meant to constitute an exhaustive set of rules but rather serve as representative examples of the pragmatic conditions necessary for rational discourse.

value (Habermas, 1996, p. 108). This idea is often captured in expressions such as "equally good for all" or "equally in everyone's interest" (Habermas, 1996, pp. 153, 255).[626] More precisely, for Habermas (U) can be derived from the basic presuppositions of argumentation and from a related but less demanding principle of Discourse (D), which states that "only those norms of action are valid to which all possibly affected persons could assent as participants in rational discourses." (Habermas, 1990, p. 66, 1996, pp. 107, 459).

Therefore, while (D) functions as a minimal condition of rational justification, (U) serves as a stronger biconditional principle, purportedly grounding the very idea of moral validity (Habermas, 1996, pp. 107–108). However, this derivation of (U) from (D) has faced criticism (Finlayson, 2019, p. 32). Scholars argue that the premises from which Habermas attempts to derive (U)—namely, the normative conditions of discourse and a weak conception of normative justification—are themselves embedded with rich moral assumptions, particularly drawn from his theory of modernization.[627] As such, the formal derivation of (U) may not hold, since it appears to presuppose the moral content it aims to establish. Habermas, however, was well aware of this potential criticism as he repeatedly expressed concern about positing outlandish metaphysical entities as the truth-makers for moral statements (Habermas, 1990, pp. 56–62, 68–76). This is why he frames his theory as a rational reconstruction of the norms and principles already implicit in social practices: he is not imposing external moral norms, but rather uncovering the "cognitive content" of morality that is already present and phenomenologically attested in everyday moral conflicts (Habermas, 1990, p. 120).

Habermas' task is practical: he starts with the constatation that mature moral agents in post-conventional societies tend to act for the sake of morality and the task of philosophy is to justify how such moral judgments are possible and rational—not to create them from scratch (Habermas,

[626] "The "oughtness" of binding norms has the absolute sense of an unconditioned and universal obligation; what "one ought to do" claims to be equally good for all." (Habermas, 1996, p. 255).

[627] Indeed, the fact that (U) contains such a demanding necessary and sufficient condition of the validity of any moral norm, so that very few moral norms survive the test of universalizability, is not so much a function of his theory but a reflection of the historical and social reality of modern society. Habermas explicitly points out that, although the number of moral norms diminishes in the course of modernization, their importance increases, citing as evidence "the growing consensus concerning human rights." (Habermas, 1993, p. 91).

1990, pp. 33–41, 1996, p. 71).[628] The derivation of (U) from (D) is not a matter of deductive proof, but rather of abductive reasoning: (U) is proposed as the most compelling philosophical reconstruction of the moral point of view suited to modern, pluralistic societies. This means that (U) is not a self-evident axiom; rather, it functions as a regulative ideal, clarifying how morality can continue to operate and construct moral judgment and action coordination in a world without "meta-social guarantees of the sacred" or a universally shared ethical framework.[629]

In metaethical terms, thus, Habermas develops a form of moral cognitivism that remains irrealist (Habermas, 1990, p. 120)[630], by arguing that moral claims are reason-sensitive and can be justified through practical argumentation. While he acknowledges that normative statements differ fundamentally from descriptive ones—the former aiming at rightness, the latter at truth—he still identifies an important analogy between truth and rightness (Habermas, 1990, p. 56),[631] because

[628] Here Habermas explicitly draws on Lawrence Kohlberg's developmental moral psychology, particularly his theory of moral stages (Habermas, 1990, pp. 33–41, 1996, p. 71). Kohlberg posited a sequence of moral development through three levels—pre-conventional, conventional, and post-conventional—each comprising two stages. The highest level, the post-conventional, includes a fifth stage of social contract reasoning and a sixth stage based on universal ethical principles such as justice and human dignity. Kohlberg maintained that stage six represented the most cognitively and morally advanced form of reasoning, as it was based on universal, self-chosen moral principles rather than merely conforming to social norms. Habermas however refines and surpass this account (Habermas, 1989). First, he disputes the notion that stages five and six are "natural" developmental stages; rather, their superiority must be established philosophically, in terms of their greater reflexivity and cognitive complexity. Second, Habermas contends that stage six moral consciousness should not be understood as an individual's application of abstract principles, but rather as a disposition to engage in dialogical procedures of moral justification. Thus, in his view, the most adequate reconstruction of mature moral reasoning is not Kohlberg's monological stage six, but the dialogical moral standpoint embodied in Principle (U).

[629] Nonetheless, Habermas's abductive approach diverges from Rawls's coherentism use of reflective equilibrium because it entails a transcendental-pragmatic stance grounded in the structure of language and communicative interaction. Habermas notice that certain norms and principles are necessarily presupposed by the very act of engaging in rational communication aimed at mutual understanding. This shift places him closer to a form of constitutivism, in which the conditions of normative discourse are not optional or externally imposed but are inescapable features of communicative practice (Heath, 2014). These discursive presuppositions function as internal criteria that participants must accept simply by virtue of engaging in argumentation. As such, they diminish the need for further external justification, since the very act of participating in discourse already commits one to these normative constraints (inescapability thesis). Nevertheless, the parallel with contemporary forms of constitutivism remains limited, for their concern is not with the communicative structure of language, but with the constitutive features of agency that directly generate the demand for justification and universality (Ferrero, 2009, p. 309). See also footnote n.660 and n.727.

[630] "All cognitivist moral theories in the tradition of Kant take into account the three aspects Kohlberg focuses on to explain the idea of the moral. The advantage of the position Apel and I defend is that the basic cognitivist, universalist, and formalist assumptions can be derived from the moral principle grounded in discourse ethics." (Habermas, 1990, p. 120).

[631] "When employing normative utterances in everyday life, we raise claims to validity that we are prepared to defend against criticism. When we discuss moral-practical questions of the form "What ought I to do?" we presuppose that the answers need not be arbitrary; we trust our ability to distinguish in principle between right norms

both are specifications of a more general underlying concept of validity (V), understood as 'rational acceptability. "(V) For any x: If x is valid, then x is amenable to rationally motivated consensus" (Finlayson, 2019, p. 36).[632] This consensus is public because the participants adopt an intersubjectively expanded perspective (first-person plural of the "we-perspective") in which they integrate each other's worldviews and self-understandings in a non-coercive and non-distorting way (Habermas, 1998, p. 58). This implies the assumption of an ideal role, in which each person actively explores their own interests while also allowing their evaluation to be shaped by the exploration of everyone else's interests (Habermas, 1996, p. 228).[633]Therefore, the agreement is not a mere contractual compromise of interests or a strategic modus vivendi, but rather a *cognitive* and *epistemic* agreement.[634]

This reconstruction should have shown how principle (U) operates as a universalizability test that is dynamic and requires actual participation in discourse with other participants.[635]

---

or commands and wrong ones. But if normative sentences do not admit of truth in the narrow sense of the word "true," that is, in the same sense in which descriptive statements can be true or false, we will have to formulate the task of explaining the meaning of "moral truth" or, if that expression is already misleading, the meaning of "normative rightness" in such a way that we are not tempted to assimilate the one type of sentence to the other. We will have to proceed on a weaker assumption, namely that normative claims to validity are analogous to truth claims." (Habermas, 1990, p. 56).

[632] As Finlayson observes, "Habermas does not say nearly enough about what the underlying generic concept of validity is." (Finlayson, 2019, p. 36). Nevertheless, Habermas attempts to reconstruct the connection between reasons and ideal justifiability in discourse through a general framework akin to principle (V), which serves as the bridging point between the two more specific principles of truth and rightness. (T): If *p* is true, then *p* is amenable to rationally motivated consensus. (R): For any norm *n*, if *n* is valid, then *n* is amenable to rationally motivated consensus.

[633] "In this speech situation, persons for and against a problematic validity claim thematize the claim and, relieved of the pressures of action and experience, adopt a hypothetical attitude in order to test with reasons, and reasons alone, whether the proponent's claim stands up. The essential intuition we connect with this practice of argumentation is characterized by the intention of winning the assent of a universal audience to a problematic proposition in a noncoercive but regulated contest for the better arguments based on the best information and reasons. It is easy to see why the discourse principle requires this kind of practice for the justification of norms and value decisions: whether norms and values could find the rationally motivated assent of all those affected can be judged only from the intersubjectively enlarged perspective of the first-person plural. This perspective integrates the perspectives of each participant's worldview and self-understanding in a manner that is neither coercive nor distorting. The practice of argumentation recommends itself for such a universalized ideal role taking practiced in common. As the reflexive form of communicative action, argumentation distinguishes itself socio-ontologically, one might say, by a complete reversibility of participant perspectives that unleashes the higher-level intersubjectivity of the deliberating collectivity. In this way, Hegel's concrete universal is sublimated into a communicative structure purified of all substantive elements." (Habermas, 1996, p. 228).

[634] Nonetheless, although it aims at a rationally grounded consensus, such consensus remains fallible and always open to further challenges and revisions should errors or misunderstandings emerge (Habermas, 1998, p. 36)

[635] This is precisely the element of reciprocity that Forst will recover and integrate with that of generality (Forst, 2012, p. 20).

Habermas see it as an "intersubjective reformulation" of Kant's Formula of Universal Law , suggesting that the moral point of view can only be realized under communicative conditions that ensure everyone tests the acceptability of a norm from different perspectives. Unlike Kant's FUL, (U) establishes the validity of a moral norm by assessing whether it satisfies the interests of each individual—not just their a priori interest in conforming to moral law. Habermas, therefore, critiques Kant's notion of universality—much like his critique of Rawls—for its monological structure, which centers on what a solitary moral agent can will as a universal law without contradiction. In contrast, Habermas emphasizes the intersubjective processes of justification that occur within a communicative community, where the validity of norms depends on their acceptability to all affected parties under the practical conditions of rational discourse.

5.1.3 Shared Strengths and Limitations of Rawls and Habermas

Their shared intellectual lineage, evident commonalities, and mutual appreciation might suggest that the Habermas–Rawls debate unfolds merely as a kind of "familial disagreement", marked by a productive exchange in which each pushes the other to refine their respective theories and carry the Kantian tradition to its fullest development (Habermas, 1995, pp. 109–110). [636] This is indeed one of the most fruitful ways to frame their dialogue (Finlayson, 2019, pp. 1–11). Yet, Habermas's critique of Rawls remains both pointed and substantial regard to the universalizability of Rawls's procedure. While Rawls grounds moral justification in the device of the original position—ultimately based on the standpoint of a solitary moral agent behind the veil of ignorance—Habermas argues that moral norms attain validity only through intersubjective discourse among real participants. In contrast to Rawls, he introduces his principle (U) as a dynamic and participatory reformulation of Kantian universalizability, one that assesses whether norms are acceptable to all affected individuals, rather than simply whether they are logically

[636] "Immanuel Kant posed the fundamental question of morality in such a way that it admitted a rational answer: we ought to do what is equally good for all persons. Without espousing Kant's transcendental philosophical background assumptions, Rawls renewed this theoretical approach with particular reference to the issue of the organization of a just society. In opposition to utilitarianism and value skepticism he proposed an intersubjectivist version of Kant's principle of autonomy: we act autonomously when we obey those laws which could be accepted by all concerned on the basis of a public use of their reason." (Habermas, 1995, p. 109).

coherent or morally generalizable in the abstract. In this way, Habermas shifts the emphasis from abstract, monological reasoning to communicative justification rooted in shared deliberation.

However, this critique has not gone unchallenged. In an article aptly titled "Why There Is No Issue Between Habermas and Rawls," Christopher McMahon forensically examines Habermas's reasons for rejecting Rawls's device of the original position as a monological, rather than dialogical, model of moral consciousness, arguing, in the end, that there is no substantive disagreement in their proposals (McMahon, 2002). McMahon shows that Habermas misrepresents Rawls's original position as purely monological, and claims that both approaches are epistemic methods aimed at identifying impartial norms; furthermore, even Habermas's discourse ethics must rely on an external notion of impartiality (like Rawls's original position) because pure argumentation cannot independently generate moral impartiality (McMahon, 2002, p. 124).[637]

However, even if one accepts that framing the distinction between the two procedures in this way—the original position as monological and the (U) principle as dialogical—is ultimately misplaced, the core issue that seems to remain is the kind of universality each one endorses. The gradual Rawls' preference for the term objectivity over universality is not merely terminological: it marks a substantive difference between the two approaches that also affects how notions of morality and publicity are employed. Rawls criticizes Habermas's discourse theory for engaging with topics and relying on theories that go "far beyond political philosophy" (Rawls, 1995, p. 135).[638] Habermas, on the other hand, denies that a political theory can—or should—avoid philosophical questions: political theory inevitably makes moral claims, and these should not be "relegated to the black box of comprehensive doctrines" (Habermas, 2011, p. 112). At the same time, Habermas argues that, despite presenting itself as constructivist, Rawls's theory remains

[637] Another scholar who has emphasized that there should be no disagreement about incorporating the same level of substantive elements found in Rawls's original position into Habermas discourse ethics is Thomas McCarthy (1994). "[Habermas] maintains that philosophically based moral theories can do no more than explicate the moral point of view, and he allows that this cannot exhaust the "semantic potential," as he puts it, of comprehensive doctrines. Precisely because of its highly formal nature, discourse ethics is compatible with different substantive conceptions of the meaning and value of life. Inasmuch as these differences figure in judgments regarding the common good, they will translate into differences on substantive questions of justice, that is, on whether specific laws or policies are equally in the interest of, or equally good for, all. Here, I think, Habermas has to move in the direction of Rawls. Indeed his emphasis on the conditions and procedures of democratic deliberation, rather than on its outcomes, might be seen as already pointing in that direction."(McCarthy, 1994, p. 57).

[638] "Indeed, the aim of his theory of communicative action is to give a general account of meaning, reference, and truth or validity both for theoretical reason and for the several forms of practical reason" (Rawls, 1995, p. 135).

fundamentally substantive rather than procedural (Habermas, 1995, p. 116, 1998, p. 72). But even if "the relation between moral discourse as procedure and the substantive results of the discourse process is very different from the relation between the original position and the principles of justice" to say that the former is strictly procedural and the latter not is "simplistic and misleading", since "in both theories, morally substantial elements are contained in the procedures themselves." (Finlayson, 2019, p. 165); the common view of Habermas's theory being more strictly procedural than Rawls's is misleading and in reality both theories blend procedural and substantive elements (Finlayson, 2019, pp. 167–168).[639]

In the end—just as noted at the beginning of this section—both philosophers rightly accuse each other of failing to fully grasp the implications of their respective theories of public reason. Habermas criticizes Rawls for proposing a form of public reason that is too limited (epistemically narrow) and not sufficiently *public* in its procedural structure (Habermas, 2011, p. 113)**,** while Rawls accuses Habermas of advocating a form of public reason that is too *comprehensive* and that implicitly prioritizes political autonomy (Rawls, 1995, p. 135). To clarify these tensions, we turn to the broader philosophical contrast between jusnaturalism and juspositivism as a lens through which to understand their disagreement.

Habermas explicitly characterizes Rawls's political theory as a form of modern jusnaturalism within the tradition of social contract theorists such as Hobbes and Locke, where the hypothetical "original position" functions analogously to the classical *state of nature* (Habermas, 1979, p. 184, 2008, p. 64). While Rawls have stressed many times that this original position is not a historical or anthropological claim, but rather a "purely hypothetical situation characterized so

[639] "In fine, there are various notions of procedure in play in Habermas's discourse theory of law and politics. In some cases, the distinction between procedure and substance is not sharp, since those procedures either contain a degree of sublimated moral and ethical substance or depend on external substantive moral and ethical resources. But the claim that Habermas's theory is more strictly procedural that Rawls's is not borne out."(Finlayson, 2019, p. 167). These considerations suggest that Habermas's criticism—that Rawls should maintain "the procedural conception of practical reason free of substantive connotations by developing it in a strictly procedural manner"—should be taken with a large grain of salt. "The comparison Habermas draws for polemical purposes is too stark and simple to do these justice. The difference between Habermas's and Rawls's theories is not, as Habermas claims in Between Facts and Norms, that Rawls's normative theory makes "contact with social reality in an unmediated way" whereas Habermas's reconstruction of the legal system does not (BFN, 65). Rather, the difference is that they offer alternative accounts of where the normative substance of constitutional democratic states lies and of how it relates to legal and democratic "procedures." Gledhill nicely sums this up as a contrast between a Rawlsian theory that embeds procedures in substance and a Habermasian one that locates substance in procedure" (Finlayson, 2019, p. 168).

as to lead to a certain conception of justice" (Rawls, 1971, pp. 17–18, 1993, p. 264 fn),[640] Habermas still critiques this formulation as overly abstract and detached from real-world legal and institutional structures; Rawls's theory is overly preoccupied with moral legitimacy while neglecting the formal and institutional aspects of law, particularly the coercive structures that underpin legal systems, leading to a "morally oriented theory of political justice," which is unrealistic (Habermas, 2008, p. 64),[641] and fails "to bridge the chasm between ideal theoretical demands and social facts" or "facticity and validity" (Habermas, 1996, p. 64).[642] This imbalance, Habermas contends, reveals a residual foundationalism in Rawl's jusnaturalism with an overreliance on moral intuition and rationalist constructions rather than on the intersubjective processes of democratic will-formation that "generates a priority of liberal rights" and "demotes the democratic process to an inferior status" (Habermas, 1995, p. 128, 2011, p. 113).[643]

Rawls counterattacks Habermas from the opposite end of the spectrum: in attempting to avoid any *substantive* moral commitments at the outset, Habermas risks stripping the concept of law of any pre-legal moral content, relying too heavily on procedure alone to generate validity (Rawls, 1995, pp. 138, 179). This leans toward a kind of procedural juspositivism that, in Rawls's

---

[640] "In justice as fairness the original position of equality corresponds to the state of nature in the traditional theory of the social contract. This original position is not, of course, thought of as an actual historical state of affairs, much less as a primitive condition of culture. It is understood as a purely hypothetical situation characterized so as to lead to a certain conception of justice." (Rawls, 1971, pp. 17–18).

[641] He argues that Rawls's theory is "developed in vacuo" (Habermas, 2008, p. 57), "first grounded in abstracto" (Habermas, 2008, p. 64), and "justified before philosophical experts" (Habermas, 2008, p. 59), rather than arising from the democratic procedures and discursive practices that, for Habermas, are essential to legitimate lawmaking.

[642] "In any case, an ethical theory does not lie at the level where it could make up for the weakness of Rawls's attempt to bridge the chasm between ideal theoretical demands and social facts. The resistant reality with which critical reason wants to keep in touch is not just, and not even primarily, made up of the pluralism of conflicting life ideals and value orientations, of competing comprehensive doctrines, but of the harder material of institutions and action systems. At the first stage of theory construction, Rawls is already concerned with questions of how one can legally implement the principles of justice that were first grounded in abstracto. And he does not fail to recognize the aspect of enforcement by which coercive law is externally coupled with its addressees' behavior-in contrast to morality, which can appeal to the sense of justice alone. But the relation between positive law and political justice stands in need of further clarification. Rawls concentrates on questions of the legitimacy of law without an explicit concern for the legal form as such and hence for the institutional dimension of a law backed by sanctions. What is specific to legal validity, the tension between facticity and validity inhabiting law itself, does not come into view." (Habermas, 1996, p. 64).

[643] "[P]ublic autonomy appears in the first instance as a means for realizing private autonomy" (Habermas, 2011, p. 113)

view, may fail to adequately safeguard individual rights: although Habermas insists on the *co-originality* of private and public autonomy—that is, of individual rights and popular sovereignty—Rawls accuses him of effectively granting *priority* to public autonomy over private autonomy (Rawls, 1995, p. 169).[644] Rawls suggests that Habermas's emphasis on "radical democracy" as a precondition for the rule of law risks subordinating individual liberties to the democratic process, even though Habermas explicitly aims to avoid either form of subordination (Habermas, 1996, pp. 122–125).[645]

Both critiques raise important points and are, in a sense, valid. Although Rawls explicitly rejects metaphysical foundationalism, Habermas rightly emphasizes that his construction of the original position and his notion of "pre-fabricated" principles of justice can be seen as an implicit attempt to ground justice on substantive foundations that precede the democratic process (Habermas, 1995, pp. 128–129). Rawls, at least in a theory of justice, does subscribes a shadowy conception of natural rights and natural duties that "hold between persons irrespective of their institutional relationships" (Rawls, 1971, pp. 114–115),[646] that might criticized as both excessively onto-foundationalist (Korsgaard, 1996b, pp. 35–36) and overly socio-formalist (O'Neill, 2016, p. 83 fn).[647]

At the same time, Rawls also correctly identifies a risk of foundationalism in Habermas when he argues that his discourse theory is "burdened by many controversial theoretical

[644] Rawls quotes a sentence in Between Facts and Norms in which Habermas claims that "the sought for internal connection between popular sovereignty and human rights lies in the normative content of the very mode of exercising political autonomy" (Habermas, 1996, p. 103). As Finlayson (2019, p. 194) noted, however, Habermas's point in that passage is that the content of the general will is shaped not only by the decisions made but also by the internal discipline of the discourse—specifically, the extent to which arguments adhere to the rules governing that discourse. Rawls's objection would be more aptly directed at Habermas's strategy of deriving underspecified basic rights as functional prerequisites of democratic deliberation.

[645] Habermas's co-originality thesis does not assert that neither private autonomy—and the rights that enable it—nor public autonomy enjoys priority; it asserts that both do. Habermas demonstrates the co-originality of private and public autonomy by showing how the basic rights are called into being by the democratic process and vice versa. In doing so, it sub-ordinates neither the subjective basic rights of citizens (and the liberties those rights enable) nor the democratic process to external constraints exerted by an antecedent moral order of rights. As Finlayson exemplified, the relation of co-originality has four main axes: equal priority, equal weight, mutual irreducibility, and complementarity or interdependence (Finlayson, 2019, pp. 98–99).

[646] "Help another when he is in need or jeopardy is a natural duty" (Rawls, 1971, p. 114).

[647] As O'Neill underlined, in Rawls "the principles of natural duty are derived from a contractarian point of view" (O'Neill, 2016, p. 83 fn).

assumptions" that extend "well beyond political philosophy" (Rawls, 1995, p. 135), similarly to a "logic in the broad Hegelian sense," one that "encompasses all the allegedly substantive elements of religious and metaphysical doctrines." (Rawls, 1993, pp. 378–379).[648] While Habermas vigorously denies to subscribes the claims of jusnaturalism, especially that the authority of legitimate law is borrowed from the authority of an antecedent moral order (Habermas, 1996, p. 314),[649] he also explicitly states that human rights have a "fully moral content" (Habermas, 1996, pp. 536–537) and that they are the legal and institutional embodiment of valid moral norms (Habermas, 1996, pp. 99–106).[650] He tries to maintain that his theory avoids jusnaturalism drifts by grounding democratic legitimacy entirely within the political domain, understood as the realm of enforceable law, and by exemplifying the "logical derivation" of a system of basic rights from the form of law and from principle (D), which is conceptually prior to the distinction between

[648] "Thus, according to Habermas, the substantial elements of Hegel's view of Sittlichkeit, an apparently metaphysical doctrine of ethical life (one among many possible examples), are-so far as they are valid-fully sublimated into (I interpret him co mean expressible, or articulated, by) the theory of communicative action with its procedural presuppositions of ideal discourse. Habermas's own doctrine, I believe, is one of logic in the 6road Hegelian sense: a philosophical analysis of the presuppositions of rational discourse (of theoretical and practical reason) which includes within itself all the allegedly substantial elements of religious and metaphysical doctrines. His logic is metaphysical in the following sense: it presents an account of what there is human beings engaged in communicative action in their life world. As to what 'substance' and 'substantial' mean, I would conjecture that Habermas intends something like the following: people often think that their basic way of doing things-their communicative action with its presuppositions of ideal discourse, or their conception of society as a fair system of cooperation between citizens who are free and equal-needs a foundation beyond itself discerned by a Platonic reason that grasps the essences, or else is rooted in metaphysical substances. In thought, we reach behind or deeper to a religious or metaphysical doctrine for a firm foundation. This reality is also expected to provide moral motivation. Without these foundations everything may seem to us to waver and we experience a kind of vertigo-a feeling of being lost without a place to stand." (Rawls, 1993, pp. 378–379).

[649] The core of the co-originality thesis is precisely that neither has priority over the other and that both are equally important (Habermas, 1996, p. 314).

[650] While this may initially appear to be another terminological disagreement (centered on the concept of morality) rather than a substantive one, there is more at stake. Habermas's distinction between ethics and morality, with the "conceptual priority of the right over the good" (Habermas, 1996, p. 286) is introduced precisely to argue that his theory does not presuppose the validity of any particular substantive ethical conceptions of the good life, since he explicitly denies that this is a possibility for a post-metaphysical theory under modern conditions. "The perspective of justice and that of evaluating one's own life are not equally valid in the sense that the morally required priority of impartiality can be leveled out and reversed in favor of the ethical priority of one's particular goals in life" (Habermas, 1996, p. 287). The only substantive ethical element in discourse theory then are "remnant of the good at the heart of the right" and the "dogmatic core" of the idea of public autonomy (Habermas, 1996, pp. 536–537). Yet at the same time, Habermas asserts that for human rights, and laws in general, to be legitimate, they must "harmonize with the universal principles of justice and solidarity" (Habermas, 1996, p. 99) and that "a legal order can be legitimate only if it does not contradict basic moral principles." By "moral principles" (moralischen Grundsätzen), here he means valid moral norms (Habermas, 1996, p. 106), so not principle (U) but the substantive norms, the validity of which is established through actual discourse among participants.

morality and law (Habermas, 1996, pp. 106–108).[651] Nonetheless, critics such as Benhabib (1986, 1992) and Larmore (1996, 1999) have highlighted the potentially foundationalist implications (more moral than pre-moral) of both principle (U) and principle (D), suggesting that Habermas's theory may not entirely escape the very metaphysical commitments it seeks to reject.[652]

Furthermore, as Rawls aptly noted, the risk in Habermas's proposal is also that in attempting to resist the "Platonic impulse" toward foundationalism, he ends up offering a formalist foundation for legitimacy and morality (Rawls, 1993, p. 379). The tension between facticity, the positive enforceable nature of law backed by state coercion, and validity, which concerns law's claim to general recognition and rationally motivated obedience, does not seem to be fully resolved simply by affirming their "co-originality" (Habermas, 1996, p. 314). This notion, particularly the co-originality of human rights and popular sovereignty, appears insufficient to fully account for the normative foundation of universal basic rights, without avoiding charges of circularity. Habermas's attempt to justify such rights exclusively within the legal-political framework, without appealing to independent moral grounds, has been criticized also by constructivists, precisely for risking a reduction of their moral substance to mere legal form, thereby overlooking their pre-legal normativity (Forst, 2012, p. 109).

---

[651] "Naturally, to conceive (D) with sufficient abstraction, it is important that we not limit a fortiori the kinds of issues and contributions and the sorts of reasons that "count" in each case. The moral principle first results when one specifies the general discourse principle for those norms that can be justified if and only if equal consideration is given to the interests of all those who are possibly involved. The principle of democracy results from a corresponding specification for those action norms that appear in legal form. Such norms can be justified by calling on pragmatic, ethical-political, and moral reasons-here justification is not restricted to moral reasons alone." Habermas emphasizes this point also explicitly in contrast to Rawls, whom he sees as placing too much weight on the normative dimension of law—its legitimacy—while overlooking its social reality or *facticity*, including its institutional form, mechanisms of enforcement, and practical implementation. "Rawls's "political" conception of justice responds to a problem that Hegel posed in terms of the relation between morality and ethical life. For the social-contract approach to natural law, the problem of the relation between norm and reality first arose at a · different level. Rational natural law, having distinguished law from morality, took into consideration the tension between facts and norms built into positive law itself. This gave it from the start a more realistic focus than a morally oriented theory of political justice. It confronted the reality of the political process, one might say, in its entire breadth. If Rawls wanted to take in this issue, he would not stop, at the second stage of his argument, with a reflection on the conditions of an accommodating political culture. Rather, he would have to get involved in a normatively informed reconstruction of the historical development of the constitutional state and its social basis." (Habermas, 1996, p. 65).

[652] Benhabib argued that (U) was redundant, and that (D) should be given a "weak justification programme" which adduced supplementary moral premises of "universal respect" and "egalitarian reciprocity" (Benhabib, 1986, p. 308, 1992, p. 31). Habermas's tendency to add stronger supplementary premises for his justification of (U) seems indeed a tacit acknowledgement of his failure to provide either a formal derivation or a "rationalist" justification of (U).

And yet, this leads to the opposite recurring question of whether such rights should instead be conceived as natural rights: basing universal rights solely on the notion of nature is problematic, as it risks reification and metaphysically burdensome assumptions.[653] Both Rawls and Habermas seek to avoid this outcome—at least to the extent that they reject grounding rights in a moral order that precedes and transcends positive law (Habermas, 1996, p. 314; Rawls, 1993, p. 45). However, in doing so, their conception of universality blends substantive and procedural elements, ultimately exposing them to the very foundationalism and formalism they aim to overcome. This raises a pressing question: how can a constructivist conception of universality offer a genuine alternative to the shortcomings of both jusnaturalism and juspositivism? Can recent developments within the constructivist framework succeed in avoiding the reintroduction of foundational assumptions in their account of universality without collapsing into formalism if they fail to anchor their procedures in some substantive conception of reason?

5.2 EMERGING DEVELOPMENTS IN CONSTRUCTIVIST UNIVERSALITY

5.2.1 O'Neill's Reasonable Universalism

The most contemporary expressions of the constructivist framework are built directly on the path laid out by Rawls and Habermas. In this context, Onora O'Neill stands out as one of the first prominent figures, who develops her thought precisely in critical dialogue with Rawls (especially in response to the coherentist drifts associated with his notion of universality). In her early works, O'Neill returns to the foundations of Kantian philosophy to offer an alternative "anti-

[653] The concept of nature might be problematic even in the reductionist-evolutionary sense, if it implies that human nature could have evolved differently—potentially without reason. While this possibility does not negate the universalist character of reason, it does cast doubt on the sufficiency of grounding rights solely in nature, especially if normativity is not understood as constitutive of human nature itself. The most effective way to refer to natural rights prior to the completion of a constructivist procedure is by appealing to the constitutive role of reason, as recent strands of constructivism suggest. This, however, is also something that Habermas, more or less explicitly, anticipates when he argues that a fully universal principle of normative validity can be grounded in the normative presuppositions inherent in speech and action—presuppositions that communicative agents necessarily adopt when they participate in practical discourse. In this view, it is the procedure of discourse itself—not a comprehensive metaphysical conception of reason or the person—that generates normative constraints.

Cartesian" reading of Kant's practical philosophy (O'Neill, 1989, pp. 3–7).[654] She interpret the Categorical Imperative, in all its formulations, as a procedure to set maxims (O'Neill, 1989, pp. 83, 129),[655] that could be adopted simultaneously by all members of a possible world of "interacting and non-self-sufficient" agents (O'Neill, 1989, pp. 141–142).[656] The central feature common to the three formulations is, thus, precisely universalizability: "any principle of thinking and acting that can have authority cannot enjoin principles on which some members of a plurality cannot (not "would not"!) act." (O'Neill, 1989, p. 20).[657]

---

[654] "On the reading that I shall propose the first Critique is not only deeply antirationalist but profoundly political." (O'Neill, 1989, p. 4)

[655] O'Neill clarifies that for Kant, a maxim is a "subjective principle of action" (O'Neill, 1989, p. 83) but is not "equated simply with intentions, which may be multiple, some of them profound and others superficial" (O'Neill, 1989, p. 129). Rather, a maxim functions as the "the underlying or fundamental principle of an action in the sense that any other principles to which the act conforms are selected and explicable because that is what it takes to act on a certain maxim in that situation (as perceived by the agent)." (O'Neill, 1989, p. 129). This uniqueness is essential: if actions had multiple maxims, the Categorical Imperative could not serve as a moral guide. The particular nature of the world means that there will always be aspects of an action that cannot be universally adopted; however, this does not make the action morally objectionable if the underlying maxim is universalizable. "A maxim, Kant tells us, is "a subjective principle of action"; it is "a principle on which the subject acts" (G, IV, 42 m; cf. 40 m). A maxim is therefore the principle of action of a particular agent at a particular time; but it need not be "subjective" in the sense that it seeks to fulfill that particular agent's desires. In speaking of maxims as subjective principles Kant is not adopting any sort of heteronomous standard, but means to propose a standard against which the principles agents propose to act on, of whatever sort, may be tested. The Categorical Imperative provides a way of testing the moral acceptability of what we propose to do. It does not aim to generate plans of action for those who have none. Although maxims are the principles of action of particular agents at particular times, one and the same principle might be adopted as a maxim by many agents at various times or by a given agent on numerous occasions. It is a corollary of Kant's conception of human freedom that we can adopt or discard maxims, including those maxims that refer to our desires." (O'Neill, 1989, pp. 83–84).

[656] "The Categorical Imperative in the FUL and FEI formulations asks us to solve a certain sort of simultaneous equation: to determine a set of maxims that could be adopted simultaneously by all members of a possible world of interacting and non-self-sufficient beings. To do this we may adopt either the perspective of one who checks whether all others can follow certain proposed guidelines for action, or the perspective of one who asks whether all who are acted upon will retain the capacities for agency that would permit action on the proposed guidelines. The two perspectives can be combined, as they are in FKE, when we ask whether the maxim(s) in question can be adopted in a "systematic union of rational beings under common objective laws - that is, a kingdom" (G, IV, 433). The thought experiment of the kingdom of ends is one where we consider ourselves both as acting (as hypothetical universal legislators) and as acted upon (as hypothetical subjects to those laws, whose agency would be destroyed by a law that made others into mere means or failed to treat them as ends-in-themselves). Each formulation of the Categorical Imperative is then a way of testing maxims for their conformity to the basic requirements of a possible community of beings who are and remain capable of action, despite the vulnerability to one another's action of their capacities to act." (O'Neill, 1989, pp. 141–142).

[657] It is important to emphasize that, for O'Neill, universalizability is not the sole central component of Kant's theory of practical reason. She repeatedly stresses that there are two essential components, both, not only to Kant's theory but to any viable account of practical reason: universalizability and instrumental rationality. "Plausibly the two central components of Kant's theory of practical reasoning — universalizability and instrumental rationality — must constitute parts of any theory of practical reasoning. Traditions ossify if they become esoteric or elitist, as traditions whose fundamental principles cannot be universally shared must do. (I take it that the fundamental principle even of

It is especially peculiar how O'Neill frames the different formulations of the Categorical Imperative as distinct perspectives on the same moral question faced by agents. Her way of clarifying the distinction between the Formula of Universal Law (FUL) and the Formula of Humanity as an End in Itself (FEI) is both textual and illuminating (O'Neill, 1989, pp. 128–129).[658] The FUL requires agents to act only on maxims they could will as universal laws, acknowledging others as agents capable of doing the same. In contrast, the FEI demands that agents avoid actions that would undermine the capacity of others to act—that is, actions that would impair their agency. Thus, while the FUL emphasizes the formal universality of moral principles and the FEI emphasizes the substantive preservation of agency, they are ultimately two sides of the same coin, since each helps define and support the other (O'Neill, 1989, pp. 142–143).[659] This mutual dependence clarifies Kant's claim that the different formulations of the Categorical Imperative are ultimately equivalent in principle, even though they emphasize different aspects of practical reasoning.

All three formulations—including and especially the Formula of the Kingdom of Ends (FKE)—exemplify Kant's conception of public as the capacity to reason about principles that

---

some sorts of aristocratic tradition can, though it need not, be shared by nonatistocrats.) Traditions fade if their fundamental principles are no longer appropriately expressed in ancillary intentions (as it seems the Tory tradition has faded, despite surviving rhetoric, after a century of living in clothing stolen from Whigs, monetarists and other liberals). (O'Neill, 1989, p. 159). For a general explanation of the role of instrumental reasoning according to various constructivist authors, see section 2.2.3

[658] "A first account of the difference between the three formulations might stress the differing perspectives from which this agent's question is taken up. FUL addresses the question from the perspective of agents who acknowledge that others too are agents and enjoins them to shun principles that could not be adopted by others, that is, that could not be universal laws. FEI addresses the agent's question from the perspective of agents who acknowledge that action affects others, and enjoins them to avoid damaging others' capacities to act. To settle whether or not FUL and FEI are equivalent, the answers the agent's question receives when explicated from these two perspectives must be compared." (O'Neill, 1989, pp. 128–129).

[659] "We might ask how we can think of the difference of perspective between FUL and FEI as a difference between giving the form and giving the matter of the Categorical Imperative. FUL focuses on maxims agents can universally act on, that is, on the possible "form" of universal agency; FEI focuses on the maxims according to which agents can treat others, who are in a sense recipients or "matter" of their action, if those others' agency, that is, ability to adopt maxims, is not to be impaired. FUL emphasizes the form that action must take if action of the same form is to be possible for all; FEI emphasizes the constraints that preserve the "matter" that makes agency possible. Just as the figure and the ground of a pattern may be mutually determining, and yet leave us in no doubt which we call the figure and which the ground, so the universalizable form of maxims of duty and the agency-respecting content of maxims of duty may be mutually determining without being indistinguishable. Indeed, only if the relation between the form and matter of the Categorical Imperative is mutually determining in this way can Kant's claim that the formulae are at bottom the same and yet distinguishable make sense. If the form and matter of maxims of duty are in this way mutually determining, Kant's other claims about the relation betwen the formulations of the Categorical Imperative are more readily understood." (O'Neill, 1989, pp. 142–143).

everyone can (and not merely "wants to" or "would") put into practice. O'Neill agrees with Habermas that, although Rawls uses the term "public reason", his approach is more Rousseauian than Kantian (O'Neill, 2000, pp. 72–73 fn 19); Rawls grounds his conception of reason in the sharing of viewpoints or debates among fellow citizens, thereby risking reducing the publicity of reason within a bounded society (O'Neill, 2001, p. 186, 2016, p. 104). At the same time, O'Neill distances herself from Habermas's conception of public reason, reaffirming that Kant roots reason not only in discourse but in the principles that sustain it.[660] Communication, as Kant sees it, is a voluntary activity among free and rational beings that poses a practical challenge: we must identify and justify principles that enable and sustain meaningful communication among diverse individuals, and Kant does so by linking the development of reason to the development of practical reason—how we communicate, tolerate, and form judgments in public life. The authority of reason does not stem from external sources, but from its capacity to enable coherent communication and self-reflection, therefore, we must ask which maxims or practical principles best guide us in our efforts to communicate, and "must try to avoid principles that could not regulate communication among a plurality of separate, free and potentially reasoning beings." (O'Neill, 1989, p. 43).[661]

The approach is dialectical, not transcendental: O'Neill begins negatively, by ruling out non-universalizable maxims of communication—such as those based on coercion or falsehood (O'Neill, 1989, pp. 44–45),[662] in order to arrive at some proactive maxims that pass the burden of

[660] This marks more clearly the emergence of the constitutivist line of argument, which also emphasizes self-reflection—a dimension that Habermas, as noted in footnote n. 629, does not fully integrate into his framework.

[661] "The division Kant makes between the natural emergence of some cognitive capacities and the dialectical development of others in human communication is fundamental to his picture. Communication is action, and hence at least to some extent freely undertaken rather than a natural product. It takes place between beings who are at least partially separate from one another, and at least partially free and rational. This picture does not deny that human communication has an animal and evolutionary basis; indeed, Kant's developmental account of reason fits well with an evolutionary view of cognition. But no noncultural account of human communication would be complete. Human communication is not a set of repertoires whose emergence reflects only the evolution of the species and the maturation of individual organisms, but has a history. Neither are the principles of communicating that emerge in the course of this history given from any source that transcends human life. They have to develop and be instituted in the course of human communication. There is neither a natural nor a preestablished harmony in the conversation of mankind. Because the structure of human communication is not preestablished, its conduct is a practical problem. We are not guaranteed coordination with others, so we must ask which maxims or practical principles can best guide us when we seek to communicate, and must try to avoid principles that could not regulate communication among a plurality of separate, free and potentially reasoning beings." (O'Neill, 1989, p. 43).

[662] "One initially plausible, but on reflection impossible, maxim of communication, which he discusses in scattered passages in the first Critique, is that of polemical debate (at other points he speaks of contention, of quarreling, of eristic dispute). Could we make the fundamental guideline of our communicating parallel to a maxim of war making, with the aim of victory? If such a maxim were fundamental, no holds would be barred. Victory would

universalization. If we find such maxims of communication, their justification must be recursive: they will be principles by which practices of communication can be maintained and developed rather than stultified, and they can be maintained as self-sustaining precisely because they emerge from within human practices rather than being imposed from outside. For this reason, in addition to the three formulations of the categorical imperative, O'Neill shows that Kant also presents three recursive maxims of public reasoning, though she refers to them using different terms such as "sensu communis", "public sense" (O'Neill, 1989, p. 45),[663] and "common human understanding" (O'Neill, 1989, p. 25).[664]

The first of these maxims is thinking for oneself, also referred to as the "maxim of unprejudiced thought" (O'Neill, 1989, p. 46) and the "maxim of a never-passive reason", which involves the autonomy of thought and the ability to form one's own judgments without relying on external authorities or prejudices (O'Neill, 1989, p. 26).[665] This is the most crucial maxim for the self-preservation of reason: for genuine communication and debate, all parties must be guided by this maxim, otherwise, understanding and agreement would be spurious—mere echoes of what another or the majority assert.

The second maxim is to think from the standpoint of everyone else, which is also described as the "maxim of enlarged thought" (O'Neill, 1989, pp. 26, 46). "Only those who try to think from

take priority over comprehension itself. But the victory that rests on incomprehension is not a victory in debate or communication. The aim of discussion or debate cannot reduce to victory; it can at best be victory by securing the other's agreement or understanding or conviction, or perhaps "agreement to differ". However, if these are the aims of discussion, communication must be guided by maxims that have some regard to other's being able to follow the communication. Achieving another's compliance in debate does not amount to winning the debate (intimations here of the master slave dialectic). A maxim of coercion in debate is no more universalizable than other maxims of coercion. Another alluring, but on reflection impossible, maxim of communication to which Kant turned his attention (in a form few of his admirers find adequate) is that of falsehood." (O'Neill, 1989, pp. 44–45).

[663] O'Neill (O'Neill, 1989, p. 45 fn 11) explains why "public sense" is the preferred translation—especially to contrast it with "private sense"—over other alternatives such as Arendt's "community sense" (Arendt, 1959, pp. 208–209, 280–284).

[664] "The Categorical Imperative, applied to reasoning itself, demands that we reason only on principles that others can (not "will" or "would") act on. To do so is to adopt three "maxims of common human understanding" (CJ, V, 294). The first of these is (1) to think for oneself." (O'Neill, 1989, p. 25).

[665] "Those who reject this discipline no longer seek to interact or communicate; they are not committed to the maintenance or the development of any sharable modes of disciplining thought or action, and so may find their supposed reasoning impenetrable to others. Kant calls this maxim the "maxim of a never-passive reason" (CJ, V, 294). To adopt it is to refuse to submit to alien forces, such as the will of others, superstitution or prejudice." (O'Neill, 1989, p. 26).

the standpoint of everyone else and strive to listen to and interpret others and to see the point of their contributions are genuinely aiming to be "fellow workers" and to avoid maxims to which others cannot agree." (O'Neill, 1989, p. 26). This formulations captures the reflexive but grounded nature of reasoning: not founded on transcendent norms, but grounded on the commitment to think in ways accessible and justifiable to others, to communicate *with* them, rather than *beyond* them, to reach "the world at large" (O'Neill, 1989, pp. 46–47).[666] It is through this second maxim that publicity and toleration starts overlapping (O'Neill, 1989, pp. 28, 50).[667]

The third maxim of public reason is to "always think consistently" (O'Neill, 1989, pp. 26, 47). This is not just a formal requirement of logic but a continuous challenge: as we revise our judgments in light of others' perspectives, inconsistencies can arise. That's why Kant calls this maxim "the hardest to attain" and an "unending philosophical task" (O'Neill, 1989, p. 47), and says it can only be achieved through the integration of the first two maxims (O'Neill, 1989, p.

[666] "Without a transcendental viewpoint, taking a universal or detached view cannot be a matter of adopting any neutral, Archimedean standpoint, but only one of seeking to see one's own initial judgments from the standpoints of others. One who adopts a maxim of enlarged thought must therefore listen to what others are actually judging and communicating. There is no lofty position above the debate, as perhaps there might be if human reason had a transcendent source. There is only the position of one who strives to reach and understand the perspectives of others, and to communicate with rather than past them. Only communication that conforms to the maxim of enlarged thought can reach "the world at large". This is what the cleric who reasons publicly in What Is Enlightenment? must do if his communication is to be fit to reach "the world at large". Practices of toleration that include respect for others and their no doubt partial and private understanding are as fundamental to communication as is self-respect." (O'Neill, 1989, pp. 46–47).

[667] In the Kantian constructivism framework, toleration is not just a political virtue but the essential condition for enabling diverse, potentially reasoning individuals to establish the full authority of reason and engage in unrestricted debate about justice and political order. "[M]any uses of reason must remain relatively private in the sense that they can in principle be used to communicate only with some restricted group, defined by another, rationally ungrounded, authority. These deprived and partial private uses of reason may sometimes help to establish more complete and public standards of reason. But there is no guarantee that they will always do so. Since they are at least partly shielded from refutation and correction by one or another nonrational authority, such reasonings may embody principles that could not survive open scrutiny and are not generally indispensable for progress toward a more comprehensive or more generally shared rationality or the possibility of a just polity. Toleration in the Kantian picture is then not merely a political virtue or a practice that would have to be part of any achieved just polity. It is the only matrix within which a plurality of potentially reasoning beings can constitute the full authority of reason and so become able to debate without restrictions what a just political constitution might be." (O'Neill, 1989, p. 50). This line of reasoning is closely akin to the one adopted by Rawls in his discussion of agreement on religious grounds. Forst (2012, 2013), in turn, develops a more articulated account of toleration, distinguishing between its *concept*—structured by the components of objection, acceptance, and rejection—and its various historical *conceptions* (he favors a "respect conception"). On this view, toleration is not merely an epistemic or pragmatic precondition for coexistence, but a moral relation among justificatory equals, according to which individuals must refrain from coercion unless the reasons for it can be reciprocally and generally justified to all those affected (Forst, 2012, p. 148).

26).[668] Once again, we see that the construction and vindication of reason is recursive—it demands constant self-correction and openness to others to point us our inconsistencies. Reason, in this view, is not a solitary faculty of consistency but one that must always be tested and shaped through communication with others (O'Neill, 1989, p. 27).[669]

The maxims of communication are, thus, not static rules but iterative principles that constrain and refine possible forms of interaction. In this way, they mirror the recursive structure of reason itself, which continually reflects upon and disciplines its own processes—not to generate fixed algorithms, but to establish constraints on thought and action that can be shared among others (O'Neill, 1989, p. 27).[670] Accordingly, the Categorical Imperative does not prescribe specific actions or beliefs; rather, it constrains what may be thought, said, or done in ways that remain open to the reason of others—ensuring that no principle adopted undermines the very conditions of a public deliberation (O'Neill, 1989, p. 47).[671]

Having this framework in mind, another central distinction for clarifying the notion of universality in O'Neill is the distinction between abstraction and idealization. Idealization involves the introduction of specific and potentially implausible conceptions of agents or contexts, which

---

[668] "The set of judgments that we independently form, then revise as we shift our standpoint to take account of others' standpoints, will constantly change, and so may repeatedly fall into inconsistency. Hence Kant claims that this maxim is "the hardest of attainment" and "only attainable by the union of both the former" (CJ, V, 295)." (O'Neill, 1989, p. 26).

[669] "Critique of reason is possible only if we think of critique as recursive and reason as constructed rather than imposed. The constraint on possibilities of construction is imposed by the fact that the principles are to be found for a plurality of possible voices or agents who share a world. Nothing has been established about principles of cognitive order for solitary beings. All that has been established for beings who share a world is that they cannot base this sharing on adopting unsharable principles. Presumably many specific conformations of cognitive and moral order are possible; in each case the task of the Categorical Imperative is not to dictate, but to constrain possibilities for acting and for cognition. Theoretical rationality constrains but does not determine what can be thought or believed, just as practical rationality constrains but does not dictate what may be done." (O'Neill, 1989, p. 27).

[670] "[Kant] does not deify reason. The only route by which we can vindicate certain ways of thinking and acting, and claim that those ways have authority, is by considering how we must discipline our thinking if we are to think or act at all. This disciplining leads us not to algorithms of reason, but to certain constraints on all thinking, communication and interaction among any plurality. In particular we are led to the principle of rejecting thought, action or communication that is guided by principles that others cannot adopt, and so to the Categorical Imperative." (O'Neill, 1989, p. 27).

[671] "The grounding of principles of reasoning in incipient communication is mirrored by the grounding of developing communication in principles of reasoning. The supreme principle of practical reason both emerges from and disciplines human communication. Its breach, whether by failure to judge for oneself or by indifference to others' communications or to consistency, damages not only particular communications, but also the practices of reasoning on which possibilities of communicating rest." (O'Neill, 1989, p. 47)

makes the reasoning inaccessible to those who do not fit the ideal (O'Neill, 2000, p. 80).[672] Abstraction, in contrast**,** is a necessary process which focuses on common characteristics (such as the finitude and interconnectedness of human beings) that political theories must undertake in order to adapt to the absence of a homogeneous community and culture: "Without (considerable) abstraction, there is no communication with those of differing culture and no account of justice that can be made accessible to diverse audiences; in short, there is nothing that is universally relevant." (O'Neill, 2016, pp. 88–89).

O'Neill, as we mentioned, criticized Rawls precisely for employing idealizations, as this limited the universal scope of his theory (O'Neill, 1989, pp. 207–209). While Rawls presents his theory of justice as grounded in abstraction—particularly through the original position and the veil of ignorance, which strip away personal characteristics such as desires and social status—O'Neill contends that this method goes beyond mere abstraction: the original position relies on a highly idealized conception of the human agent, one that assumes individuals are mutually independent and motivated solely by instrumental reasoning aimed at securing primary goods (O'Neill, 1989, p. 209).[673] Such idealization, though theoretically appealing, is misleading; it might produce elegant models, but it overlooks the embodied, vulnerable, and fundamentally relational nature of human beings (O'Neill, 1989, p. 209).[674]

---

[672] "Much recent Kantian work has indeed been predicated on unvindicated idealizations, which undermine its applicability to human life. Kant's approach may be read in another sense, in which the finitude of human beings, of human rationality and the connected ness among human beings, is stressed rather than denied. If this reading of Kant's work can be sustained, the charges made by its contemporary opponents will, at the very least, need revision. If it cannot, then the challenge of constructing an account of ethics, and specifically of justice, whose premisses rely only on genuine abstraction without idealization remains open rather than shown impossible" (O'Neill, 2000, p. 80).

[673] "[I]s there more than abstraction behind the construction of the original position? At first it may seem that Rawls does no more than abstract. The veil of ignorance merely obscures; it only limits claims about agents in the original position. They simply know less than actual human agents. Rawls does not add to their information or desires in any respect. However, the ways in which he abstracts are governed by a certain ideal. The artful tailoring of the veil of ignorance is determined by a highly selective abstraction from actual human choosing, which reflects a certain ideal of the human subject. This ideal is used not just to determine how "thick" or "thin" the veil of ignorance should be, but rather to tear some carefully placed holes in it. Abstraction, taken strictly, is unavoidable and in itself innocuous. Idealization is another matter." (O'Neill, 1989, p. 209).

[674] "The veil of ignorance described in A Theory of Justice was tailored to hide the interlocking structure of desires and attitudes that is typical of human agents. Once the social relations between agents were masked, it could seem plausible to assign to each desires for a uniform shortlist of primary goods, and to build a determinate ideal of mutual independence into a conception of justice. This ideal is not met by any human agents. It is not only deficient and backward human agents whose choosing would be misrepresented by these ideal agents of construction. The construction assumes a mutual independence of persons and their desires that is false of all human beings. Such independence is as much an idealization of human social relations as an assumption of generalized altruism would

By contrast, O'Neill argues that genuine universality is not grounded in the idealizations of detached human agents, but in the abstraction of shared features of human agency (O'Neill, 1989, pp. 209–210).[675] The difficulty with this notion of universality, she suggests, does not lie in its normative assumptions, since some degree of normativity is inevitable even in more moderate form of "reasonable empiricism" (O'Neill, 1989, p. 66),[676] but rather in a common misinterpretation: namely, the belief that universality entails rigid, uniform treatment that ignores the particularities and complexities of individual cases (O'Neill, 2000, p. 68).[677] O'Neill challenges

be." (O'Neill, 1989, p. 209). In later works, primary goods are no longer understood by Rawls as means to whatever ends individuals might pursue, but as necessary conditions for the development and exercise of the "moral powers"—the capacity for a sense of justice and a conception of the good (Rawls, 1980, 1993). These capacities come to define the highest-order interests of persons and are built into the design of the original position. Consequently, Rawls's theory becomes explicitly dependent on a normative ideal of the person—autonomous and rational—reflecting a procedural Kantianism that avoids metaphysical commitments. This is still not enough to avoid some "unkantian" idealization of the person, like is independent nature (O'Neill, 1989, pp. 210–211).

[675] "Idealizations have no doubt many theoretical advantages: above all, they allow us to construct models that can readily be manipulated. However, they may fail to apply to any significant domain of human choosing. This failure is sometimes defended by construing ideals not as abstractions but as "simplifications." This too is strictly inaccurate. A theory simplifies if it either leaves things out (i.e., abstracts) or smooths out variations. If it incorporates predicates that are not even approximately true of the agents to whom the model is supposed to apply, it does not simplify. If idealizations do not "simplify" the descriptions that are true of actual agents, then they are not innocuous ways of extending the scope of reasoning. They covertly assume and endorse "enhanced" versions of specific human characteristics and capacities. Idealization masquerading as abstraction produces theories that may appear to apply widely, but in fact covertly exclude from their scope those who do not match a certain ideal. They privilege certain sorts of human agent and life by presenting their specific characteristics as universal ideals." (O'Neill, 1989, pp. 209–210).

[676] "Although few have been as explicit as Rawls in stating the objectives of contemporary Kantian work on justice, I believe that nearly all of its exponents have taken from him the aim of reaching more or less Kantian conclusions from more or less empiricist views of reason, action and freedom. If the enterprise is to succeed, the empiricist conceptions which are to form the starting point must be free of the types of difficulties which Rawls and many others have seen in Kant's starting points. If these are the ambitions of contemporary Kantian work on justice, it is rather surprising that it has constantly been targeted by critics of many differing persuasions with the very criticisms that they level against Kant." (O'Neill, 2000, p. 66).

[677] "Evidently it is not easy to avoid idealizations. The critics of contemporary Kantian work on justice may have a point (which they misleadingly speak of as an objection to abstraction) in that 'the canons of a reasonable empiricism' too may harbour idealized conceptions of reason, freedom or action. I shall return to this point. The second group of criticisms levelled against Kant and Kantian work charges that any position which argues for universal principles renders it ethically deficient. Universality may be understood in various ways. Minimally, a principle is formally universal if it applies to all cases within its scope – whether that scope be large or small. More ambitiously, a principle may be said to be universal in scope if its domain is very large – for example, cosmopolitan. However, the feature that has generally worried critics of Kant and of Kantian ethics is neither of these sorts of universality, but rather the thought that any (universal) principle mandates uniform treatment, thereby overlooking all the differences between cases, and requiring inflexible similarity of treatment. However, if universal principles are abstract they will underdetermine action, and so will not prescribe uniform action. Neither universal form nor wide scope entails uniform requirement. It is not therefore in the least surprising that many of the universal principles advocated by Kant and by proponents of Kantian conceptions of justice patently do not require uniform treatment." (O'Neill, 2000, p. 68).

this view by distinguishing between the form and scope of universal principles and their action-guiding content. Abstract universality—understood as a principle's applicability across relevant cases—does not require identical prescriptions in every situation; on the contrary, such principles typically underdetermine action, allowing for differentiated and context-sensitive application (O'Neill, 1989, p. 224).

By distancing herself from Rawls's conception of universality, O'Neill seeks to articulate a distinct constructivist account by focusing on "modal questions"—that is, inquiries into the possible choices available to abstract (but not idealized) agents (O'Neill, 1988, p. 17, 1989, p. 218).[678] These are *modal* rather than *hypothetical* questions, concerned not with what agents would choose under imagined conditions, but with what they could—or could not—consistently will as rational beings. The modal question to start the practical inquiry is: 'Which principles must a plurality of abstractly characterized agents reject?' This question might be answered in a negative, yet still action-guiding way, namely, by rejecting non-universalizable principles—those whose adoption would destroy, impair, or undermine the agency of others. Paradigmatic examples include, as we already mentioned, principles grounded in coercion, deception, or violence, all of which directly compromise the ability of others to act (O'Neill, 1989, pp. 215–216, 228, 1996, p. 50, 2016, p. 59).[679] Within a world populated by rational yet vulnerable agents, such principles fail

[678] "Several features distance the alternative constructivism sketched here from Rawls' pioneering project. First, idealised accounts of agents, their rationality and their mutual independence are explicitly rejected: there is no appeal, however oblique, to transcendent moral claims. Second, abstraction from the determinate desires of agents is complete: hence no special ingenuity is needed to avoid either heteronomy or cruder forms of relativism. Third, and as a consequence, it is impossible to answer questions about the hypothetical choices of abstract agents. Fourth, the construction therefore has to fall back on modal questions about the possible choices of abstract agents, and construct an answer to the question 'What principles must a plurality of abstractly characterised agents reject?'. Finally, rejection of non-universalizable principles can guide action by requiring that we ensure that the agents actually affected, with their particular identities and vulnerabilities, can genuinely choose or refuse those principles. If this sketch can be filled out, there is at least some space between realism and relativism" (O'Neill, 1988, p. 17, 1989, p. 218).

[679] "Examples of principles that cannot be universalized can illustrate the point. A principle of deception, which undermines trust, would, if universally adopted, destroy all trust and so make all projects of deception impossible. Selective deception is not improbable; universal deception is. A principle of coercion, whose enactment destroys or undercuts the agency and willing of at least some others for at least some time, cannot be universally followed. Those who are at a given time the victims of coercion cannot act, so cannot make coercion their own principle. Equally, action on a principle of violence damages the agency of some, so cannot be universally acted on. Put quite generally, principles of action that hinge on victimizing some, whether by destroying, paralyzing or undercutting their capacities for agency for at least some time and in some ways, can be adopted by some but cannot be universally adopted. (O'Neill, 1989, pp. 215–216). She brings up these examples many times across the same book and in the subsequent works: "If we are looking to work out a set of universal obligations, principles of action that cannot be universally adopted must be rejected. One type of principle that can be adopted by one agent yet cannot be adopted by all is a principle whose enactment destroys or disables or undercuts agency in at least some others on at least some occasions and so renders those others at least temporarily unable to adopt that same principle. For example,

the test of universalizability. And yet, this negative approach simultaneously leads us toward principles that are universalizable—such as the principle of charity, understood in its classical sense as concern for the needs of others (O'Neill, 1989, pp. 229–230).[680] Charity becomes a duty because egoistic indifference would undermine the agency of others and could not coherently be willed as a universal law.

To summarize the argument thus far, O'Neill offers a compelling clarification of the concepts of universalizability and universality in a constructivist vein. Universality is the characteristic or property of principles and claims that describes whether something applies to all individuals or to all relevant cases (O'Neill, 1989, pp. 85–88); a principle has a universal form if it applies to all cases within its scope (O'Neill, 2000, p. 68).[681] Universalizability, instead, is the

---

principles of coercion or deception or violence undercut or destroy somebody's agency in some respect; so their rejection is part of any set of principles that can hold universally. What is going on here is a strict version of Kantian universalizing. It identifies as principles of obligation those principles that must be adopted if nonuniversalizable principles are rejected." (O'Neill, 1989, p. 228). Another interesting and detailed example that she gives many times is that of a principle of injuring severely: "Nobody can coherently think of a principle of injuring severely as adoptable by all (in any domain): if it were so adopted, some would succeed in acting on it, hence other would become their victims, so unable to act on it. There is no possible world in which adoption of a principle of severe injury by all could be achieved; there is nothing agents could do that would count as or contribute to bringing about such a world." (O'Neill, 1996, p. 50).

[680] "Here we are considering agents who are embodied in vulnerable form, are subject to illness and death and need regular nourishment and at times in their lives the care and attention of others. They need these things not merely for some arbitrary purpose but as conditions of becoming and remaining agents. Among such rational and needy beings, a principle of self-centered indifference to others' needs could not be universally adopted. Its attempted adoption would erode or destroy the capacities for agency of some of their number, thereby guaranteeing that its adoption would be less than universal. If rational and needy beings are to act only on principles they all can act on, they must eschew principles of undercutting or destroying one another's capacities to adopt principles. They must make it a matter of principle not merely not to coerce or deceive, but to reject self-centered indifference to others' needs insofar as this indifference threatens the other's agency. This means that they must make it a matter of principle to help meet at least some of the needs of at least some others. Among rational and needy beings, charity, in the old sense of concern for the needs of others, is an obligation."(O'Neill, 1989, pp. 229–230).

[681] O'Neill distinguishes between the *form* of universality—its internal coherence and the possibility of being willed as a universal law—and its *scope*—the external extension of its domain of application. Principles "can combine universal form with restricted scope" (O'Neill, 2016, p. 2), for example a principle willed by all the members of a specific society; conversely, a principle may aspire to a very broad and "cosmopolitan scope" (applying to everyone) while remaining formulated in non-universal terms (O'Neill, 2000, p. 68). This shows that neither, taken individually nor together, requires uniform treatment, precisely because universal principles are abstract and must be contextualized through practical judgment. "Universality may be understood in various ways. Minimally, a principle is formally universal if it applies to all cases within its scope – whether that scope be large or small. More ambitiously, a principle may be said to be universal in scope if its domain is very large – for example, cosmopolitan. However, the feature that has generally worried critics of Kant and of Kantian ethics is neither of these sorts of universality, but rather the thought that any (universal) principle mandates uniform treatment, thereby overlooking all the differences between cases, and requiring inflexible similarity of treatment. However, if universal principles are abstract they will underdetermine action, and so will not prescribe uniform action. Neither universal form nor wide scope entails uniform

process of testing or verifying whether a maxim is suitable to become a universal law (O'Neill, 1989, p. 89); it focuses on the coherence and capacity of a maxim to be adopted by all without self-contradiction or negating the possibility of action (it is an internal criterion of the action itself). In other words, a principle is universalizable if it can be coherently willed as a universal law, and a universal law is universal if it applies to all relevant cases. The test of universalizability is the means by which one determines whether a principle can possess the property of universality and thus, in the practical field, be morally valid or just. In clarifying this distinction, O'Neill carefully separates universalizability from other test of universalization, like the Golden Rule based on what "everyone desires" or "would want" (O'Neill, 1996, p. 125 fn).[682] Such approaches, which ground moral acceptability in desire or inclination, are considered heteronomous and are rejected by both Kant and O'Neill: universalizability concerns what can be coherently willed for all, not what is or would be commonly shared (O'Neill, 1989, pp. 131–132).[683] O'Neill rigorous reinterpretation of universality is able to provide a characterization of principles that counter the longstanding charge that Kantian ethics collapse into empty formalism (O'Neill, 1989, p. 158).[684]

---

requirement." (O'Neill, 2000, p. 68). Furthermore, as we will later see, for O'Neill "the most fundamental feature of human rights is their scope: whatever else they are, human rights are standards for everyone."(O'Neill, 2016, p. 126).

[682] "Hence universalizability is distinct both from formal universality and cosmopolitan scope, from universal prescriptivism and from conceptions of ethical universalism, such as those stated in Golden Rules, which advocate action on principle that would be accepted or would be desired by all. Universalizability is a spare, modal notion and incorporates no claims about the content or ethical significance of motivation or preference. It merely demands that principles adopted be ones that could be principles for all in the relevant domains" (O'Neill, 1996, p. 125 fn).

[683] "FUL states that we should act only on those maxims through which we can will at the same time that they be universal laws. This is often misconstrued as a claim that morally worthy maxims must be ones that we are willing, that is, want, to see universally adopted (cf. Golden Rules, Universal Prescriptivism). This may not be such deeply dyed heteronomy as a utilitarian pursuit of maximal satisfaction of desires; but it is heteronomy nonetheless, and Kant rejects it decisively (e.g., G, IV, 43cm). Kant's understanding of FUL is uncompromisingly rationalist. He asks whether we can without contradiction will (not "want") a maxim (underlying principle) to hold as a universal law." (O'Neill, 1989, pp. 131–132).

[684] "These arguments to show that some maxims cannot consistently be universalized are sufficient to show that the charge of empty formalism, which has so often been made against Kant, does not hold. The nonuniversalizability of fundamental principles is clearly of deep moral significance. In general, claims about the moral status of more specific intentions and acts holds only for determinate contexts. Only if some more specific principles enjoin actions that are indispensable in all circumstances for action on a morally worthy maxim can specific principles of action themselves be shown to be morally worthy. But such is the variety of human ways of life that few actions can be held indispensable in this way. Even the avoidance of lying (at least in a narrow verbal sense) may be dispensable in a nondeceiving Trappist community. The avoidance of theft and fraud may be dispensable means to avoiding deception in a society in which property institutions are still lacking. There are no specific intentions that are indispensable for all who adopt a maxim of not always refusing help or the development of abilities. This suggests again that moral rules, because they prescribe and proscribe specific types of action, can indeed not be derived without substantive premises about specific sorts of social arrangements: Since the means that are indispensable for action on a morally worthy maxim are typically contingent on social arrangements, Kant's picture of right action is inadequate.

This crucial insight is translated by O'Neill in the political global scale: in an increasingly interconnected world—characterized by the "porous borders" created through the global movement of goods, capital, and people—a universalist approach to justice becomes imperative (O'Neill, 2000, p. 201).[685] O'Neill argues precisely that principles of justice are fundamentally cosmopolitan rather than civic; their applicability is not conditioned by membership in a political community but by the existence of moral agents capable of acting on principle. Justice, therefore, cannot be confined to national or communal contexts: if we recognize others as agents with whom we interact, even across vast distances, we cannot justifiably exclude them from the scope of justice (O'Neill, 2000, pp. 187–188, 195–198).

O'Neill criticizes conceptions of justice that while ostensibly committed to universal principles, in practice restrict their application to the confines of the state: while states may create international institutions (like the UN or WHO), these typically serve state-centered interests, and justice remains primarily focused on domestic self-interest rather than global responsibility (O'Neill, 2016, p. 218).[686] These approaches takes the front matter backwards: "Why should the

---

For in abstraction from context it is indeterminate what the outward aspects of morally worthy action are, and hence indeterminate to what (merely) right action should conform. Rechtslehre not merely cannot be part of a critique of practical reason, it cannot be part of a metaphysic of morals that establishes "legal" duties for the human condition. It can be part only of more local determinations of rules that must be followed within particular social and cultural settings. This does not, however, require us to adopt a relativist account of right action, since the moral worthiness of particular institutional settings and social practices, in terms of which judgments of right and wrong may be established, is itself open to judgment in terms of the underlying maxims by which a life shaped by such institutions and practices is informed. Traditions too may be judged morally, as indeed MacIntyre judges the liberal tradition, by their deepest principles" (O'Neill, 1989, p. 158).

[685] O'Neill's (O'Neill, 2016, pp. 177–178) is cosmopolitan in orientation, but does not advocate for a world government without borders, like Habermas (1998) does. Instead, she envisions a global order marked by "porous boundaries and dispersed power". "Over the last fifty years boundaries have been becoming more porous. It has happened gradually, selectively and in faltering ways, and has transformed both political and economic life for countless people. For example, most boundaries in Europe stand where they stood at the end of the Second World War, but are unimaginably more porous. Boundaries that were once impervious to capital flows and trade, to cultural and technical influence and above all to movements of people have become highly, if variably, porous." (O'Neill, 2000, p. 201).

[686] "If states are the primary global actors, they may be able to secure some forms of inter-national justice, but 'domestic' justice will remain the primary focus and the primary locus of any legitimate use of force. Political philosophy that starts with states is therefore unlikely to offer a full account of global or cosmopolitan justice, in which any boundaries have to be justified rather than presupposed. Any political philosophy that assigns basic moral standing to societies (or to other bounded entities such as nations, peoples or states) will be hard pressed to justify a more cosmopolitan conception of justice, including health justice. As we know all too well (in this respect realists in international relations are not wrong), states and other bounded entities usually pursue self-interest in constructing and supporting inter-national institutions." (O'Neill, 2016, p. 218).

boundaries of states be viewed as presuppositions of justice rather than as institutions whose justice is to be assessed?" (O'Neill, 2000, p. 4). She is particularly concerned with contemporary reformulations of liberal political theory, such as those advanced by Rawls in *The Law of Peoples*, which retain universalist language but restrict their practical force to relations among "fellow citizens" within "bounded societies." (O'Neill, 2000, p. 4).[687]

Given the frequent failure of states—especially weak or unjust ones—to ensure global justice, O'Neill advocates for a more robust conception of the plurality of justice agents (O'Neill, 2016, pp. 177–178). She argues that non-state actors, such as transnational corporations (TNCs) and non-governmental organizations (NGOs), also bear obligations in promoting justice (O'Neill, 2016, pp. 170–175, 188–192) For example, TNCs can insist on decent labor or environmental standards that go beyond the minimums set by local law, even when operating in weak states; even if they cannot render state more just, they can still stop them from becoming more unjust or completely dysfunctional (O'Neill, 2001, p. 194).[688]

O'Neill critique of the 1948 Universal Declaration of Human Rights must be understood precisely in this light. Any proclamation of rights will be ineffective unless the corresponding obligations are assigned to specific, identifiable agents and agencies capable of fulfilling them. While universal liberty rights (e.g., freedom of speech, freedom from torture) do indeed require universal obligations in the sense that everyone has the duty to respect them (O'Neill, 2000, pp.

[687] "These and similar criticisms levelled against ethical universalism have received considerable attention in more recent work on justice. Once again Rawls's work is paradigmatic. Although he has not repudiated universalist aspirations, and has not accepted the communitarian view that embedded social norms form adequate starting points for ethical and political reasoning, Rawls has argued in his later work that the agreements of fellow-citizens have a fundamental status in an adequate conception of the reasonable, and thereby in political justification. Many others also take it that the discourse or debate of citizens is fundamental to justice. The thought may seem convincing if we take for granted that an account of justice may presuppose that we are fellow citizens of some state, or (as Rawls puts it) of a 'bounded society', insiders who can share a common debate about justice. But the approach is strangely silent about the predicaments of outsiders, and about the justice of a world that is segmented into states, a world in which for each of us most others are emphatically not fellow-citizens. It seems to me that, on the contrary, an adequate account of justice has to take seriously the often harsh realities of exclusion, whether from citizenship of all states or from citizenship in the more powerful and more prosperous states." (O'Neill, 2000, p. 4).

[688] "Justice has to be built by a diversity of agents and agencies that possess and lack varying ranges of capabilities, and that can contribute to justice – or to injustice – in more diverse ways than is generally acknowledged in those approaches that have built on supposedly realist, but in fact highly ideologised, views of the supposed motivation of potential agents of justice." (O'Neill, 2001, p. 194).

101–102),[689] rights to goods and services (such as food, healthcare, or education) cannot be guaranteed simply by assigning identical obligations to all agents: here it is crucial to allocate obligations to specific agents and institutions. The problem with the state-centric approach is precisely that the obligations of justice tied to universal rights are assigned to institutions that, by their very nature, differentiate those rights based on nationality. Therefore, while the Declaration proclaims universal rights, it is insufficiently transparent about the correlatives obligations, which are the "active components" of justice (O'Neill, 2001, p. 185),[690] and entrusts their protection to states, thereby legitimately differentiating individuals' rights according to state borders (O'Neill, 2000, p. 180).[691]

Against these state-centric positions, O'Neill defends what she calls a "reasonable form of universalism for ethics and politics" (O'Neill, 2000, p. 3),[692] which diverges sharply from both

[689] "All advocates of rights are agreed that if there are universal claim rights to liberty, then the corollary obligations must also be universal. For example, a right not to be raped, or a right to compete with others for employment, will be marred if it is a right against some but not against all others. If there were some others who have no obligation to refrain from raping, or no obligation to let others compete in the market place, then nobody would have an unrestricted right of either sort. Mere liberties apart, universal liberty rights require corresponding universal obligations." (O'Neill, 2000, pp. 101–102).

[690] "The statism of the Declaration should not surprise us. Its preamble addresses member states who "have pledged themselves to achieve, in co-operation with the United Nations, the promotion of universal respect for and observance of human rights and fundamental freedoms." Yet since states cannot implement justice, let alone global justice, without constructing and coordinating many other agents and agencies, it is a matter for deep regret that the Declaration is so opaque about allocating the obligations of justice. The reason for regret is that in the end obligations rather than rights are the active aspects of justice: a proclamation of rights will be indeterminate and ineffective unless obligations to respect and secure those rights are assigned to specific, identifiable agents and agencies which are able to discharge those obligations . (O'Neill, 2001, p. 185).

[691] "This, it seems, is the role envisaged for states in the Universal Declaration of Human Rights. The supposed universality of the rights proclaimed in the Declaration is hedged with the claim that for any individual these rights may legitimately be differentiated at state boundaries. The Declaration proclaims that everyone has the right to a nationality (evidently intended not as (a sense of) national identity, but quite minimally as membership of a state) (Article 15), and differentiates the freedom of movement that each is to enjoy in his or her own state (Article 13) from the much more restricted rights of asylum which he or she may enjoy in certain circumstances in other states (Article 14). A basic feature of a just international order – interstatal order – as outlined in the Declaration is that any given individual's rights to travel, take up abode, work, own property, and vote and their corresponding duties may justly be circumscribed by the boundaries of the state of which he or she is citizen, or more broadly a member. The Universal Declaration does not endorse very much institutional cosmopolitanism." (O'Neill, 2000, p. 180).

[692] "Perhaps this does not matter. Universalism in ethics and politics has acquired a bad name among a wide range of philosophers, for reasons that are discussed in many of these essays. My own view, however, is that if we are to have an account of justice that is relevant for a world in which state boundaries are increasingly porous to movements of goods, capital, ideas and people, and in which state sovereignty is increasingly circumscribed, we shall need to work on setting out a reasonable form of universalism for ethics and politics. This will not be easy because conceptions of justice which were devised with the thought that states are the primary context of justice may need a lot of stretching and remodelling if they are to do global duty. Questions about the scope of ethical and political

contemporary forms of liberalism (O'Neill, 1989, p. 223) and new forms of communitarianism (O'Neill, 1989, p. 222).[693] This form of reasonable universalism leads to a form of "institutional cosmopolitanism" (O'Neill, 2000, p. 201), that is pragmatic and feasible, not utopian or idealized: it thus also avoids both naïve idealism (which might overlook the complexity of the real world) and defeatist relativism (which retreats into local frameworks abandoning universality). It acknowledges the fragility and limitations of human agency, as well as the practical difficulties of designing institutions capable of meeting global obligations; nonetheless, it still insists that ethical and political reasoning must aspire to universality, in the sense that its principles must be accessible, intelligible, and justifiable to all agents, regardless of location, citizenship, or cultural background.

In summary, O'Neill's constructivist conception of universality develops from a robust interpretation of Kantian notion of universalizability—focused on the consistency of maxims and respect for the limited and fragile nature of human agency—towards a broader political application concerning global justice. This evolution entails a growing recognition of the need for principles that transcend state boundaries and cultural particularities, a stronger emphasis on obligations along rights, and the inclusion of a variety of actors (not just states) as responsible for the realization of justice. Throughout, O'Neill avoids unsupported idealizations and maintains a strong commitment to practical feasibility.

5.2.2 Forst's Reflexive Universality

Rainer Forst, in his articulation of Kantian constructivism, integrate Rawls conception of justice (Forst, 2024, pp. 145–159, 263–268)[694] with Habermas discourse theory (Forst, 2024, pp.

reasoning and about the boundaries of just institutions cannot be treated as mere afterthoughts in an account of justice that is convincing and useful in the contemporary world." (O'Neill, 2000, p. 3).

[693] "Recent neo-Hegelian works depict abstract liberal accounts of justice and rights as expressive of a social order that eroded the communities that made sense of particular configurations of virtues. They insist that obligations can best be embodied in characters and traditions rather than in formal institutions. Justice need not be abandoned, but must be domesticated. Justice too is specific to actual communities. International justice and human rights cannot be vindicated — except within circles where these have become the prevailing outlook. The diverging fates of justice and charity in the modern period are symptoms" (O'Neill, 1989, p. 222).

[694] "Let us start from a reflection on the grammar of justice, taking our lead from Rawls, who defined the core concept at the center of different conceptions of justice as implying that institutions are just when "no arbitrary

210–211, 268–272)[695], affirming that his justificatory approach is both *recursive*, a procedure of self-legitimization without assuming any fixed content or values, and *discursive*, a practice that involve a dialogue among free and equal individuals (Forst, 2024, pp. 199–200).[696] This process of justification relies on a form of universality that Forst, from his earliest work, describes firstly as "reflective" (Forst, 2002, p. 271) and then "reflexive" or "self-critical" (Forst, 2024, pp. 62–63). By distinguishing between a "reflective" from a "subsumptive" universality (Forst, 2002, p. 271),[697] Forst is underlying that universal principles or norms are derived and justified through a process of rational deliberation or construction, rather than being passively discovered or revealed: moral norms are not simply imposed from above or deduced from pre-existing general principles, but emerge through a continuous and critical dialogue among participants, who must be willing to

distinctions are made between persons in the assigning of basic rights and duties and when the rules determine a proper balance between competing claims to the advantages of social life." (Forst, 2024, p. 157). While Forst was a direct student of Rawls, his conception of universality owes much also to O'Neill's recursive interpretation, a point he openly acknowledges on multiple occasions (Forst, 2024, pp. 8, 114, 158, 181, 200). The "overarching reflexive principle of justice" is a "recursive argument" (Forst, 2014, p. 114) not only because the principle of justification explains justification while also affirming itself, but also because justification is an evolving process in which no claim is exempt from critical examination and questioning. In this sense, you cannot escape the demand for justification, even when the principle of justification is the subject of scrutiny.

[695] Forst aligns with Habermas's critique of Rawls, arguing that Rawls inadequately addresses moral autonomy, thereby diluting the validity of universal principles of justice (Forst, 2012, p. 88). In this context, Forst emphasizes the importance of reclaiming the interconnectedness of universality with the discursive criteria of reciprocity and generality. In his words: "it is necessary to prove that the normative claim to validity really can be upheld reciprocally (i.e., without some of the addressees claiming certain privileges over others and without one's own needs or interests being projected onto others) and generally (i.e., without excluding the objections of anyone affected). Reciprocity and generality-understood here as universal are thus the decisive criteria of justification in the moral context; the former underscores the equal status of and imperative of concrete respect for moral persons as individuals, the latter prevents the exclusion of those possibly affected and confers the authority of the moral community on the individual. In this context, the general principle of reasonable justification is thus to be conceived as the principle of reciprocal and universal justification." (Forst, 2012, p. 20).

[696] "I call this form of justification moral constructivism. It is both a discursive and a recursive enterprise; discursive, because justification needs to be a practice between free and equal persons, even if it is only counterfactually possible, and recursive, because no content or value is given except that of the agents and criteria of construction. This is what the autonomy of morality really means: its substance is constituted by principles and agents of practical reason. In other words, we can only find out what is morally justified by reconstructing the principles and agents of moral justification and using these principles in the right way." (Forst, 2024, pp. 199–200).

[697] "A crime against humanity is a (moral) "crime against humankind"; it means failing to carry out the moral duty to recognize every human being in his or her moral integrity as a representative of humanity as such and to regard him or her as an authority to which one must justify oneself. To judge morally means regarding, as a human being, other human beings as members of the comprehensive community of all human beings and behaving toward each one here and now in a way that can be justified with general reasons. That is the meaning of "reflective" judgment, which does not subsume the particular under the general but equips the particular with the authority to demand general justified reasons for modes of action" (Forst, 2002, p. 271).

review and modify their positions in light of new arguments and evidence. Forst emphasizes that this process of deliberation is essential to ensure that the norms are genuinely universal, as they must be justified to all individuals involved in a transparent and impartial way.

By distinguishing recursive from discursive in his notion of reflexive universality, Forst acknowledges that he employs the notion of recursivity in a slightly different sense from O'Neill (Forst, 2012, p. 293 fn 10), stressing more the formal and monological elements of it (Forst, 2012, p. 107).[698] Reason is recursive in that one cannot escape the demand for justification, even when the very principle or procedure of justification becomes the object of scrutiny. This recursive structure means that any attempt to justify the framework of justification itself must still meet its own standards. Such recursivity not only secures the principle's self-legitimating character, but also prevents reasoning from becoming static or dogmatic. It frames reasoning as an open-ended, dynamic process in which no claim is immune from further questioning and potential revision; [699]

[698] Forst would not call it the monological element, but the self-referential structure of justification, since he wishes to avoid any possible monological implications. He follows Habermas in arguing that Rawls reconstructs the Kantian notion of public reason in a predominantly monological way, in a way that it "becomes more a theory of the 'private use of reason with a political-public intent' than a theory of the deliberative democratic use of reason" (Forst, 2012, p. 107). As noted earlier (section 5.1.3), reconstructing Rawls's proposal as monological is itself problematic (Finlayson, 2019, pp. 114, 151–152). Forst tends to do so since he wishes to associate the idea of publicity primarily to the dialogical and discursive dimensions of justification, exemplified by the criteria of reciprocity, while the element of recursivity is exemplified by the criteria of generality through the monological aspect of consistency. This emphasis might give the impression of downplaying the normative character of public reason as being both recursive and genuinely public. Yet, the divergence appears to be primarily terminological—centering on whether the notion of "public" should explicitly include the recursive dimension—rather than reflecting a substantive disagreement. For Forst, reciprocity and generality, as well as recursivity and publicity, are two sides of the same coin. There are several passages where he even treats reflexivity and recursivity as synonymous: "The principle of justification follows from a reflexive and recursive consideration of the validity claims of norms of justice that claim reciprocal and general validity for all members of a normative order. Viewed in this light, the conception of justice as justification is based only on those principles and criteria that are intrinsic to the claim to justification of justice. At the same time, the reflection in question is transcendental in nature, for it reconstructs the conditions of the possibility of rationally justified normative validity. The criteria of reciprocity and generality, understood discursively, mean that no one may make claims that they deny to others (reciprocity of content), and that the justification of such claims must be made in a language that does not accord priority to one of the perspectives (e.g., a religious one) from the outset or impose a reasonably rejectable interpretation of "true" interests or needs on others (reciprocity of reasons). Generality means that no one who is subjected to a normative order may be excluded from the discourse of justification."(Forst, 2024, p. 181).

[699] For Forst, the right of justification is excluded from this process precisely because it works as a "*fundamentum inconcussum*", even within a "postmetaphysical" Kantian constructivist framework (Forst, 2012, p. 5). Some strand of constructivism, particularly those closer with pragmatism, will argue that no claim, however foundational, including the very principle of justification itself, is immune from further questioning and potential revision (Richardson, 2013). There are some passages in Forst, especially in the early writings, that also seems to concede this, for example here: "Only by way of a recursive reflection the premises of the justification of practically reasonable action in normative contexts are reconstructed, and there is no claim to an ultimate foundation" here. But in order to close the gap (between the cognitive insight in the principle of argumentation "U" on the one side and the obligation by discursively justified norms on the other) that arises because of the Habermasian distinction […] one

their continued universality does not stem from immunity to critique but from their ongoing provisional vindication in the absence of decisive reasons for their abandonment. The recursivity of reason is, in short, regarded as a "critical and subversive force" that must be able to challenge its own "pathologies" of false practices and contents of justification (Forst, 2007, p. 302).[700]

What seems to be missing in this predominantly monological reconstruction of recursivity is the essential dialogical and reciprocal element. In a reflexive conception, however, this component is also acknowledged: justification is not merely a matter of internally revising one's own judgments in response to external stimuli but involves an ongoing, reciprocal exchange of reasons among interlocutors who hold each other accountable. This reciprocal exchange is what grounds the universal validity of claims: something become universally justifiable precisely because they withstand—and are shaped by—critique and negotiation within a community of reasoners. It is through this shared discursive process that the normative force of justification emerges, as universal claims must be accessible, intelligible, and acceptable to all affected parties (Forst, 2012, p. 146).[701]

In contrast to O'Neill, this conception of public reason is defined by Forst as a "transcendental reflection on the conditions necessary for justice claims to be valid"—namely, their capacity to be justified to all those affected; nonetheless, non-arbitrariness, is not given in advance but emerges through the actual processes of reciprocal and shared justification (Forst,

---

must emphasize the practical insight into the principle of justification and the duty and right to justification. This insight characterizes persons who understand themselves as beings who give moral reasons and act accordingly." (Forst, 1999, p. 58).

[700] "People in very different historical circumstances, speaking different 'thick' normative languages, questioned given justifications for the structures and norms they were supposed to live under, demanding other, better reasons. The struggle for justifications is, to use one of Honneth's own phrases, a deep 'grammar' of social conflict and emancipatory movements, leaving the form in which this demand is phrased up to them and historical circumstance. The practice of justification then is to be seen as a basic social and dynamic practice with the inherent potential of opening up conventional and exclusionary forms of discourse. Reason, as I see it, is a critical and subversive force – but reason it must be in order to be able to criticize its own 'pathologies' of false practices and contents of justification. A critical theory is in need of a free-standing as well as situated conception of practical reason." (Forst, 2007, p. 302).

[701] "Generality, then, means that the reasons for such norms need to be shareable among all persons affected, not just dominant parties." (Forst, 2012, p. 146). Here we can see that the two sub-principles constituting the principle of justification—the principle of generality and the principle of reciprocity—are inseparably connected. Each presupposes and reinforces the other, and both inherently contain recursive and dialogical dimensions.

2024, pp. 158–159).[702] The use of the "transcendental element" by Forst is to reaffirm that the starting point for identifying the conditions that make any claim to justice universally valid is the principle of justification; this transcendentality must therefore be understood purely in epistemic terms, and it does not override but rather informs the deliberative procedure in which individuals exchange reasons as both "reason-giving and reason-receiving entity" (Forst, 2024, pp. 199–200).[703]

This refined notion of universality allows Forst to critique both Habermas and Rawls for inadequately grounding universal rights (Forst, 2012, p. 109).[704] Habermas's attempt to justify rights through the "co-originality" of human rights and popular sovereignty (Habermas, 1996, pp.

[702] "At this point, we must take a reflexive turn and work our way up from the core concept of justice to a conception of justice as containing a practice of public justification, while taking care to avoid arbitrariness. If we want to overcome arbitrary social and political relations and institutions and also exclude arbitrary justifications for such relations and institutions, and if we have no "natural" or objective candidate for what "non-arbitrary" means, then we must take the principle of justification, as a principle of reason (defined as the faculty of justification), as the core of the conception of justice – call it a conception of justice as justification. According to the principle of reasonable justification, those justifications for social relations and institutions are free from arbitrariness that can withstand the discursive test of reciprocal and general justification among free and equal persons (as members of a justification community defined in a non-arbitrary way – a problem to which I will return). We arrive at the principle of reciprocal and general justification by a reflexive and recursive consideration of the validity claim of social and political justice norms that claim to be reciprocally and generally binding on all those who are part of a normative order; that is, an order that determines the basic standing within a social structural framework and is the proper context of claims to justification. Hence a conception of justice as justification relies on just those principles that are implicit in the very claim to justifiability that characterizes justice norms. We can call this reflection a transcendental one, because it reconstructs the conditions of the validity of claims and norms of justice. If we aim at non-arbitrariness as a conceptual core of our notion of justice, such a reflection is what we need to hold onto." (Forst, 2024, pp. 158–159).

[703] The process of justifying these reasons is precisely what he intends with Kantian moral constructivism. "I express the idea of the basic respect for autonomous persons in the language of a moral "right to justification" because, as pointed out above, according to the Kantian conception there is a categorically undeniable subjective claim to be respected as a normative authority who is free and equal to all others – and such a claim I call a moral right, one which is binding on every other moral person. It is not a right to some good or a right based on some interest; rather, it is based on being a person as a reason-giving and reason-receiving entity and an authority in the space of reasons. It is therefore the ground of all further claims to moral respect and to the validity of more specific moral norms, which must be justified by appeal to reciprocally and generally non-rejectable reasons. I call this form of justification moral constructivism. It is both a discursive and a recursive enterprise; discursive, because justification needs to be a practice between free and equal persons, even if it is only counterfactually possible, and recursive, because no content or value is given except that of the agents and criteria of construction. This is what the autonomy of morality really means: its substance is constituted by principles and agents of practical reason. In other words, we can only find out what is morally justified by reconstructing the principles and agents of moral justification and using these principles in the right way." (Forst, 2024, pp. 199–200).

[704] "Although Rawls does not account in detail for the core of Habermas's theory of co-originality […] he challenges the theory with important questions that make it necessary to find a position mediating between the two. In my view, a conception of the co-originality of basic rights and popular sovereignty should be based neither on Habermas's interpenetration thesis, nor on Rawls's original position. Rather, it must rest on an appropriate interpretation of the right to justification in various normative contexts." (Forst, 2012, p. 109).

122–125),[705] risks to "carry out the justification of basic rights in a manner that is overly immanent to law" (Forst, 2012, p. 109), reducing their moral content to mere legal form, thereby neglecting their pre-legal normativity.[706] Conversely, Forst criticizes Rawls's original position because it undermines popular sovereignty by giving normative priority to pre-determined principles of justice over democratic will-formation. In doing so, the Rawlsian framework effectively transforms popular sovereignty into "an organ for executing principles", rather than what it ought to be it, namely "the determining form for the active construction and constitution of the political and social basic structure" (Forst, 2012, p. 110).[707]

---

[705] Habermas argued that the internal relation between private and public autonomy requires a set of abstract rights that citizens must recognize if they want to regulate their life together by means of legitimate positive law. This "system of rights," which each concrete democratic regime must appropriately elaborate and specify, delineates the general necessary conditions for institutionalizing democratic processes of dis course in law and politics. These rights fall into five broad categories: "1) Basic rights that result from the politically autonomous development of the right to the greatest possible measure of equal individual liberties. 2) Basic rights that result from the politically autonomous development of the status of a member in a voluntary association of consociates under law. 3) Basic rights that result immediately from the actionability of rights and from the politically autonomous development of legal protection. 4) Basic rights to equal opportunities to participate in the processes of opinion and will formation in which citizens exercise their political autonomy and through which they generate legitimate law. 5) Basic rights to the provision of living conditions that are socially, technologically, and ecologically safeguarded, insofar as the current circumstances make this necessary if citizens are to have equal opportunities to use the civil rights listed in (1) through (4)." (Habermas, 1996, pp. 122–123). The first three are the basic negative liberties, membership rights, and due-process rights that together guaran tee individual freedom of choice, and thus private autonomy. The fourth, rights of political participation, guarantees public autonomy. Habermas argues that each side is indispensable and cannot simply be reduced to the other: without the first three sets of rights, there is no private autonomy (and thus no free and equal subjects of law) but without the fourth set the laws and rights guaranteeing private autonomy are merely paternalistic impositions rather than expressions of self-governance. Rights of political participation, that is, enable citizens themselves to shape and further define the rights they enjoy as "privately autonomous" and thus to become "the authors of the laws to which they are subject as addressees." Finally, a fifth category of social-welfare rights becomes necessary insofar as the effective exercise of civil and political rights depends on certain social and material conditions, for example, that citizens can meet their basic material needs.

[706] Finlayson also noted this Forst's criticism to Habermas's thin theory of the universality of basic rights. "Habermas has come in for criticism from various quarters for offering a merely legal-functionalist justification of basic rights that supposedly does not draw on moral premises. It is widely argued that his justification is normatively too thin and that he should have rested his justification more explicitly on some normative basis, whether democratic political values (Joshua Cohen), moral rights we owe to each other (Rainer Forst), a core morality consisting in a principle of equal respect (Charles Larmore), or the preconditions of the practice of moral discourse (Stefan Rummens)." (Finlayson, 2019, p. 97).

[707] Rawls rejects the accusation that he subordinates democracy to liberal rights, and he directly addresses this charge in the way it was formulated by Habermas in *Political Liberalism*: "In this section, I begin my reply to Habermas's objection […] The objections concern the correct relation among two familiar classes of basic rights and liberties-the so-called liberties of the ancients and the liberties of the mod erns. Habermas agrees that I share Jean-Jacques Rousseau's and Kant's hopes of deriving both kinds of rights from the same root. This is shown in the fact that both kinds of liberties appear equally in the first principle of justice, which is adopted in the original position. That these liberties have the same root means, he thinks, that the liberties of the moderns cannot be imposed as external constraints on the political process of citizens' self determination. He then refers to the two-stage (Zweistufig) character (127-28) of the political conception of justice as fairness, by which I take him to mean that this conception starts

In response to these shortcomings, Forst proposes a synthesizing alternative that aims to balance the moral independence of basic rights with the democratic autonomy of political self-determination (Forst, 2010, p. 725).[708] The distinction between different contexts of justice is helpful here, as it allows us to differentiate between legal autonomy—understood as the liberty of individuals to act as "accountable actors in different social spheres regulated and stabilized by law"—from the moral dimension that concerns "the content of the natural law protection of human dignity," and the political dimension that reflects "the content of positivist counter-theory that law must be the positive law of a community" (Forst, 2002, p. 264).[709] There is, thus, no contraposition between jusnaturalism and juspositivism: both are essential spheres of the validity of law, with natural law retrieving its validity through reciprocal and general arguments, and positive law retrieving its validity through consultation and lawmaking. At the foundation of both, and as a common ground across different contexts of normative justification, lies the notion of dignity conceived as a status that recognizes an entity as worthy of justification (Forst, 2024, p. 161).[710]

---

with the hypothetical situation of the original position where principles of justice are selected once and for all by the parties as equals subject to the veil of ignorance, and next moves to citizens' regular application of those same principles under the actual conditions of political life. The two-stage character of the political conception leads, he believes, to the liberal rights of the moderns having a priori features that demote the democratic process to an inferior status (127 -28). This last statement I wish to deny." (Rawls, 1993, p. 396). Rawls, in sum, affirms that both the "liberties of the moderns" (private rights) and the "liberties of the ancients" (political rights of participation) appear equally within his first principle of justice, as adopted in the original position, and he denies that the liberties of the moderns are "imposed as external constraints" on the actual political democratic processes; on the contrary, he argues that his theory reveals an "internal connection between public and private autonomy", and that "both are co-original" (Rawls, 1993, p. 432). This rebuttal, however, remains insufficient for Forst, who argues that the structure of Rawls's theory still grants priority to liberal rights over democracy, thereby limiting the actual political autonomy of citizens (Forst, 2024, pp. 263–268).

[708] "I hold that the normative grammar of human rights, both historically and systematically, calls for a conception of basic rights to democratic participation. Not just from the internal perspective of those who fought or fight for such rights, but also from a principled perspective based on the right to justification, human rights should not be seen as rights to goods necessary for the "good life." Rather, they should be seen as rights that put an end to political oppression and the imposition of a social status which deprives one of one's freedom and of access to the social means necessary to being a person of equal standing. Grounding human rights on the right to justification captures this political and social meaning of human rights as standing in opposition to earlier, as well as modern, forms of social exclusion. What is primarily at stake in inclusion is being regarded as an agent worthy of effective political justification, of giving and receiving reasons in the political realm." (Forst, 2010, p. 725). Here Forst is explicitly taking distances from instrumental approaches to human rights, especially that of Griffin (2008).

[709] Finally, law has also "an ethical dimension not because it presupposes and embodies a priori certain values, but because citizens' ethical claims, needs, and interests enter the procedure of law making, insofar as they can be justified with general reasons and can rely on general norms." (Forst, 2002, p. 266).

[710] "[I]t is important to keep in mind that a conception of political and social justice is not a form of ethical thought that needs to be "applied," because it is an imperative of political autonomy and justice that those who are subject to a normative order should become its collective authors. The authority of justice is theirs – but in the form of a collective project of emancipation bound to principles of justification and equal respect. […] The Kantian

This status entails not only a fundamental moral right to justification in all relevant interactions (Forst, 2024, pp. 199–200),[711] but also a legal and political right to participate in justificatory discourse within political contexts. "Individual rights are therefore reciprocally nonrejectable rights to respect for personal integrity. In this nonrejectability lies their moral core, which is justified neither ethically nor in terms of natural law but intersubjectively; the core must be legally determined and institutionalized." (Forst, 2002, p. 233).

Furthermore, Forst, similarly to O'Neill, extends his conception of universality at the transnational level (Forst, 2012, p. 251),[712] which he conceives as an alternative to purely *statist* visions (where justice is limited to relations between states) and *globalist* approaches (where justice is applied only at the cosmopolitan but individual level). The fundamental right to reciprocal and general justification serves to advocate for justifiable social and political relations "within and across states" and within "multiple orders of subjection" operating globally—including economic systems (Forst, 2001, p. 170).[713] It is not misleading to speak of a

---

groundwork I have been laying down is based on a fundamental moral claim of free and equal persons to be respected as autonomous normative authorities when it comes to the normative orders to which they are subject. This is my version of the Kantian idea of respecting others as "ends in themselves." Their "dignity" means that they are such justificatory authorities, and it implies a basic moral right to justification. It is a right to justification with respect to all morally relevant actions in moral contexts of interaction, and in contexts of social and political justice it is a right to participate fully in justificatory discourses about the normative order to which you are subject, so that in such contexts it becomes a legal and political right." (Forst, 2024, p. 161).

[711] "I express the idea of the basic respect for autonomous persons in the language of a moral "right to justification" because, as pointed out above, according to the Kantian conception there is a categorically undeniable subjective claim to be respected as a normative authority who is free and equal to all others – and such a claim I call a moral right, one which is binding on every other moral person. (Forst, 2024, pp. 199–200).

[712] "The first question that has to be addressed when one thinks about issues of justice that transcend the normative boundaries of states is whether one is looking for principles of international or of global justice. Whereas the former view takes political communities organized into states to be the main agents of justice (i.e., who is asked to be just and who receives just treatment), the latter takes persons, regardless of their political membership, as the primary focus of justice (at least as far as the question is concerned with who receives just treatment). On the first view, principles of international justice are to regulate the relations be tween states in a fair way; on the second view, they are to regulate the relations between all human beings in the world and to ensure their individual well-being. I shall refer to proponents of the first view as statists and those of the second as globalists. These labels are, of course, artificial and comprise a number of quite different perspectives." (Forst, 2012, p. 251).

[713] "Speaking very generally, a "transnational" approach differs from a globalist view in considering particular political contexts as contexts of justice in their own right and in constructing principles of justice for the establishment of just relations between autonomous political communities. It differs from statist views by starting from a universal individual right and by considering the global context as an essential context of justice. Given the central aim of the realization of the right to justification within and between states in order to end the vicious circle of internal and external domination, a theory of transnational justice has to combine the various contexts of justice in the right" (Forst, 2001, p. 170).

single "structure of justification" or a "justified basic structure", as local, national, and global contexts are intertwined as sites of injustice—and consequently, of justice. "There can be no global justice without internal justice, and viceversa." (Forst, 2012, p. 250).

By extending the demand for justification across borders and institutional levels, Forst is able to criticize the imposition of "Western" norms as universal (false universalism) while also opposing essentialist perspectives that marginalize voices within non-Western societies.[714] He argues that any critical theory of justice must begin with the recognition of existing injustices and acknowledges that neither the imperialistic misuse of Enlightenment ideals nor a purely local focus can address the struggles for justice embedded in the globalized modernity (Forst, 2024, p. 165).[715] The only viable approach is a self-critical universalism, which combines Enlightenment ideals of equality with a commitment to justification, to reach a *justificatory equality* (Forst, 2024, pp. 63, 266). This means recognizing every individual as a normative authority—someone who has the right to question and critique the norms and systems that govern them. This approach empowers the excluded, granting them a right to justification and critique, and avoids the pitfalls of both parochialism (exclusion) and relativism (overlooking domination).[716]

---

[714] Here is where Forst take more distance from Rawls. As he puts it: "It is a truism that any notion of justice that applies to transnational contexts needs to be properly universalizable, a requirement that raises the Rawlsian bar for theories based on ethically and culturally particular "comprehensive doctrines" to a higher level than that envisaged by Rawls himself in his reflections on liberal social pluralism in Political Liberalism. But we should also aim at more than what Rawls suggests in The Law of Peoples, namely at more than a conception of international justice from a liberal standpoint that "tolerates" non-liberal but "decent" peoples and essentially provides the "ideals and principles of the foreign policy of a reasonably just liberal people." Rather, we should come to develop a non-parochial approach, that is, one that avoids liberal as well as non-liberal one-sidedness and reified culturalist conceptions of "peoples"." (Forst, 2024, p. 154).

[715] "Reflexive universalism, as I call it, avoids such one-dimensional reductions and essentialist forms of cultural positivism by providing a normative yardstick that consists in the very idea and structure of questioning any form of domination or normative imposition. The principle of critique, of challenging normative orders, is at the core of such questioning, and thus any concrete narrative or structure of justification can be made the subject of radical questioning, which asks whether the structure in question is truly reciprocally and generally justifiable. The principle of asking for better justifications and of asking for structures of justification in the first place is as immanent to normative orders, as it transcends their historical and social forms if they are repressive, narrow, exclusionary or in other ways deficient. Since the principle leaves the authority to determine justice with those who participate in discourses among free and equal persons and thus aim to establish such discursive structures in the first place, it never speaks a language of justice that is alien to them. Rather, it locates the power to define that language with the participants and their discursive constructions." (Forst, 2024, p. 165).

[716] "Whereas some might generously interpret "the normative horizon of Enlightenment modernity" as taking "openness to criticism and reflexivity as normative goals," being willing "to unlearn," others are simply happy parochialists, i.e., neo-colonial, imperialist modern fascists who want to make an empire great again and interpret Enlightenment modernity as licensing that. What do we say to them – that they got modernity wrong? But we just learned from Allen that modernity harbors a lot of imperialist racism and fascism. So modernistic contextualism is a

To summarize, the conception of reflexive universality is central in Forst proposal to develop a critical theory of justice that navigates between the extremes of juspositivistic relationism and non-contextual jusnaturalism. As Forst explains, this approach begins "from an account of existing structures of rule and/or domination in various normative orders, ranging from the national to the international and the global, and thus develops a critical form of relationism based on a general moral right to justification. In concrete social contexts, this translates into the right to be an equal justificatory authority within the normative orders to which one is subject – either as a subject of justified rule or of domination. Thus, there is a single normative ground of justice but a range of different contexts of justice according to the nature of the subjection in question (Forst, 2024, p. 172).[717] Reflexive universality respects the diversity of contexts while upholding the demand for justification as a shared human need, which determines both a right and a correlative duty to contrast domination (Forst, 2024, p. 171).[718] This notion of universality reminds us that while our contexts may differ, the demand for justification is the one language of justice we all must speak, and that the struggle for (universal) human rights is nothing less than the struggle against domination in all its forms.

---

no-go unless it is countered by the universalism of justificatory equality, according to which each person has a justified claim to the same status of being a normative authority. In short, if false universalism based on parochialism is the problem, enlightened particularism is not the solution. The only solution is enlightened, self-critical universalism, since it is the only one that equips the excluded with a right to justification and critique."(Forst, 2024, p. 63)

[717] In the footnote, Forst clarifies that when he says that "there is only one ground of justice but various contexts of justice marked by different forms of domination" (Forst, 2024, p. 172), he knows that his use of the term "contexts" here differs from its usage in *Contexts of Justice*. This intentional shift in terminology, already employed in *The Right to Justification*, aims to underscore an alternative perspective to the conventional dichotomy between global and local contexts commonly found in debates between statists and globalists. "The main issue in this debate is to what extent the world as a whole is a context of justice, that is, a context characterized by conflicting claims that call for adjudication in light of principles of justice. For such a context to exist, there have to be identifiable authors and addressees of legitimate claims to justice, whether they are rights claims or claims based on other grounds of justice. According to the globalists, the global context is the primary context of justice, and other, more local contexts can only be legitimate once the first one is well ordered. To be sure, the statists do not deny that there are relevant claims to justice in the international sphere. They merely argue for restricting their scope and hold that an updated version of the traditional ius gentium suffices (which may entail some components of economic justice). The basic argument for this restriction is that, with respect to political and distributive justice, the globe is not the primary context of justice. Compared to the "thick" context of domestic justice, it is merely a secondary, "thin" one." (Forst, 2012, pp. 251–252).

[718] "Every human being has a right to justification and must not be subjected to domination, and every other human being in a position to help has a natural duty to do so – in addition to the more concrete relational duties of those who benefit from or uphold a system of domination – though in complex political structures the conclusion about the right course of action is very difficult to draw and any intervention needs to observe strict criteria of justifiability." (Forst, 2024, p. 171).

5.3. POTENTIALS AND LIMITS OF CONSTRUCTIVIST UNIVERSALITY

We are now in a position to grasp the distinctive character of the overall constructivist notion of universality. At its core lies the conviction that universal principles are not discovered as independent or objective facts (A. Müller, 2020, p. 6),[719] but are instead constructed through deliberation, drawing on shared procedures of reasoning accessible to all rational agents (Korsgaard, 1996b, p. 35).[720] What makes a principle "universal" is not its self-evident metaphysical status but its constitutive recursive consistency and the public capacity to withstand an intersubjective scrutiny in a forum of equals.

Yet, some critics have sought to highlight that—even though constructivists scholars agree that practical reason, through well-defined procedures, can construct universal norms—they nevertheless employ distinct procedures; for example, it has been suggested, that there is an almost linear increase in the intersubjective dimension, progressing from Kant—where it is largely absent—through Rawls, Habermas, and O'Neill, culminating in Forst. As shown in the previous pages, this reading is misleading: although there are some textual evidences to describe the Habermasian approach as more deliberative than the Rawlsian, neither Rawls's nor Kant's approach can be reduced to a purely monological procedure.

In all their procedures, universality is not an imposed uniformity of action (O'Neill, 1989, p. 54 fn 5),[721] but rather a criterion for identifying principles that can govern interactions among plural, free, and rational beings—because these principles are based on reasons that can be publicly

[719] "To be a constructivist about a certain domain is to take the facts of that domain to be an upshot, rather than the measure, of correct judgement in that domain. A constructivist about practical reasons thus denies that, when an agent comes to realize that something is a reason for a certain action, she does so in response to the corresponding normative fact, that is, a fact about what is a reason for doing what. Instead, such facts are somehow explicable in terms of our reason judgements. That does not mean that the facts are whatever we think they are. Constructivists do not deny that we can be mistaken about what is a reason and what is not. But they do suggest that normative facts are ultimately grounded in our mental states or activities." (A. Müller, 2020, p. 6).

[720] "Procedural moral realism is the view that there are answers to moral questions; that is, that there are right and wrong ways to answer them. Substantive moral realism is the view that there are answers to moral questions because there are moral facts or truths, which those questions ask about" (Korsgaard, 1996b, p. 35).

[721] "[U]niversalizable principles demand uniform action only for like cases. Conditional universalizable principles demand action only of those who meet the condition." (O'Neill, 1989, p. 54 fn 5).

and mutually justified and accepted by all, grounded in processes that ensure impartiality (Rawls, 1971, p. 190), solidarity (Habermas, 1998, p. xxxv) reciprocity (O'Neill, 2016, p. 104), and generality (Forst, 2002, p. 39). While each of these authors place greater emphasis on certain elements over others, they all acknowledge the importance of them all.

But can this procedure truly overcome the difficulties we have identified in both jusnaturalism and juspositivism? The distinctive strength of constructivism to offer a genuine alternative in the human rights debate lies precisely in its ability to chart a path between these two stagnant poles. To avoid the metaphysical queerness of jusnaturalism while at the same time resisting the moral emptiness of juspositivism, we need to ground human rights in a procedure that avoids the aforementioned pitfalls.

5.3.1 Is Constructivist Universality Foundational or Formal? Inescapability of Agency

As we have seen, one of the main critiques directed at Rawls's theory of justice is that, in avoiding metaphysical foundations, it becomes overly abstract—namely, purely formal. Critics such as Alasdair MacIntyre and Charles Taylor have argued that concepts like justice, rights, and moral reasoning must be grounded in a shared vision of the good life. Rawls, by contrast, deliberately avoids substantive moral commitments in order to preserve pluralism, focusing instead on the fairness of procedures rather than on any particular conception of the good. These communitarian critics contend that reducing justice to a formal decision-making process risks severing it from the concrete moral and cultural contexts in which real human lives are lived. Forst (2002), while still criticizing and moving beyond Rawls's notion of universality, has offered strong arguments on why these communitarian critiques were ultimately misplaced. Nevertheless, it is important to note that even these communitarian critics of Rawls—despite their differing perspectives—agreed that moving beyond both foundationalism and formalism represented a crucial and necessary step in political theory, marking a significant evolution in the debate. Their main contention lies in the belief that Rawls failed in this attempt, and that his solution led to an outcome that was not only undesirable, but, although paradoxically, undemocratic (Forst, 2002, p.

159).[722] Does the more refined version of constructivist universality succeed in overcoming this tension between foundationalism and formalism?

If the contraposition between foundationalism and formalism—and, by extension, between jusnaturalism and juspositivism—can be framed by grounding human rights either in an external metaphysical source (as in the former) or in an internal institutional code (as in the latter), then we have to prove that the constructivist notion of universality can overcome the problems associated with both approaches while preserving their respective strengths. As discussed (especially in section 2.2), constructivism manages to retain both the external and internal dimensions through the concept of the constitutive norm, which has become a central theme in the constitutivist turn of constructivist theory after Rawls.[723] Constitutivism has emerged, over the past two decades, as a significant current in debates on practical rationality, precisely because it promises a way out of the broader dichotomy that has shaped twentieth-century metaethics and the theory of action in the analytic tradition—namely, between "different forms of subjectivism and non-cognitivism, and, on the other hand, different forms of realism or objectivism" (Haase & Mayr, 2019, p. 95).[724] To overcome this impasse, constitutivism holds that the norms of practical rationality and morality possess a special kind of authority and are categorically binding for agents regardless of their contingent motives, preferences, or intentions—unlike the rules of particular games or institutions, which are binding only conditionally on one's decision to participate. In short, the argument can

[722] "In Walzer's view, Rawls's theory is an example of an undemocratic theory of democracy: it conceives a theory of justice for a political community in a abstract situation of rational choice based on certain capabilities of abstract persons and on certain abstract primary goods, which are to be distributed justly." (Forst 2002, 159).

[723] While traces of this idea appear in Rawls's work, including one of his earliest articles (Rawls, 1955, p. 5), it remains only partially developed. Nevertheless, it is entirely possible to interpret Rawls through a constructivist lens—especially, and importantly for our purposes, when he discusses natural rights and natural duties. For instance, when he states that "Help another when he is in need or jeopardy is a natural duty" (Rawls, 1971, p. 114), we may read this as pointing to the constitutive dimension of reason, or what he calls the "sense of justice." Yet, as we have already emphasized, what remain problematic of the concept of nature is the risk of reifying it, dangerously returning to the metaphor of discovery that constructivism wanted to avoid precisely by introducing the metaphor of construction.

[724] "The recent rise of constitutivism is mainly due to its promising a way out of an apparent dilemma that is widely thought to have beset twentieth century meta-ethics and the theory of practical rationality in the analytic tradition. This debate has – put in slightly simplified terms – been widely perceived as a stand-off between, on the one hand, different forms of subjectivism and non-cognitivism, and, on the other hand, different forms of realism or objectivism that posited moral norms in a special realm independent of us (as human beings or rational agents)." (Haase & Mayr, 2019, p. 95).

be analytically reconstructed as follows: 1) The norms and requirements of practical rationality and morality can be derived from the constitutive features of agency; 2) A systematic failure to comply with these requirements amounts to a loss of agency; 3) Since, in a certain sense, we cannot avoid being agents, it follows that we are necessarily bound by the duties of rationality and morality (Ferrero, 2009, p. 304).[725]

The success of this argumentative strategy depends both on demonstrating the major premise (that the norms of rationality and morality can be derived from the constitutive features of agency) and on recognizing a further implicit or unstated premise in the conclusion—namely, that we cannot avoid being agents, or in other words, that agency is not optional and "inescapable" (Velleman, 2000b, p. 113).[726] While a full examination of the extensive debate is beyond the present scope, I will simply note that there are two principal reasons for finding the thesis of the inescapability of agency persuasive: the Largest Jurisdiction argument and the Closure under Reflective Rational Assessment argument (Ferrero, 2009, pp. 308–309).[727]

The Largest Jurisdiction refers to the maximal scope and reach of agency: agency is the overarching enterprise that encompasses all others. Any engagement in a specific activity—be it a game, a social practice, or an institution—ipso facto involves engaging in the broader enterprise of agency. For example, playing chess necessarily involves exercising agency (making decisions, evaluating options, and acting intentionally) but there are forms of conduct that fall exclusively under the domain of agency and not under any other specific undertaking. A notable example is

[725] "In a rough outline, their basic claim is that the norms and requirements of practical rationality and morality can be derived from the constitutive features of agency. Hence, a systematic failure to be guided by these requirements amounts to a loss of agency. But there is a sense in which we cannot but be agents. It follows that we are necessarily bound by the oughts of rationality and morality, we are bound by them sans phrase." (Ferrero, 2009, p. 304).

[726] For Velleman, commitment to action is inescapable because we have a general disposition and a "caring" to act; this general disposition is epistemic in nature, as we are interested in understanding ourselves. Asking oneself why self-understanding should be a goal already reveals that one holds such a goal. (Velleman, p. 137)

[727] "To sum up, agency is special because of two distinctive features. First, agency is not the only game in town, but it is the biggest possible one. In addition to instances of bare intentional agency, any engagement in an ordinary enterprise is ipso facto an engagement in the enterprise of agency. Second, agency is closed under rational reflection. It is closed under the self-directed application of its distinctive discursive operation, the asking for and the giving of reasons for action. The combination of these features is what makes agency inescapable. This is the kind of nonoptionality that supports the viability of constitutivism." (Ferrero, 2009, p. 309).

"intentional transitions": the deliberate entering into, or exiting from, a particular undertaking (e.g., deciding to begin or cease playing chess). These are not "moves" within the activity itself, but nonetheless acts of agency (Ferrero, 2009, p. 308).[728]

The Closure under Reflective Rational Assessment refer instead to the fact that agency is "self-contained" with respect to rational reflection and evaluation. Unlike ordinary undertakings, from which one can withdraw in order to examine their justification (e.g., pausing a game of chess to consider whether it is worth playing at all), agency cannot be suspended in this way. Even the decision to reject agency would be an exercise of agency, thus one cannot genuinely "step outside" it. As long as one holds beliefs, entertains desires, and acts intentionally, one is exercising agency. In this sense, asking "Why be an agent?" is analogous to asking "Why exist as a self?": the very standpoint from which the question is posed presupposes the agency it attempts to evaluate (Ferrero, 2009, p. 309).[729]

The combination of these two features—its supreme jurisdiction and its reflective closure—is what makes agency inescapable. Any act of practical reasoning or reflection about the justification of agency is, by its nature, already an exercise of intentional agency. When an agent considers whether they are justified in engaging in agency, they are still operating according to the constitutive standards of agency itself. This inescapability constitutes the necessary non-optionality that underpins the validity of constitutivism, distinguishing it from other forms of formalism. The universality of practical reason is not obtained solely from the internal coherence of specific reasoning but from norms that are constitutive of the rational nature of every agent. These norms—like impartiality, generality, or the independence of evaluators—are sometimes

[728] Ferrero terms such cases "bare intentional agency". "To engage in any ordinary enterprise is ipso facto to engage in the enterprise of agency. In addition, there are instances of behavior that fall under no other enterprise but agency. First, intentional transitions in and out of particular enterprises might not count as moves within those enterprises, but they are still instances of intentional agency, of bare intentional agency, so to say. Second, agency is the locus where we adjudicate the merits and demerits of participating in any ordinary enterprise. Reasoning whether to participate in a particular enterprise is often conducted outside of that enterprise, even while one is otherwise engaged in it. Practical reflection is a manifestation of full-fledged intentional agency but it does not necessarily belong to any other specific enterprise." (Ferrero, 2009, p. 308).

[729] "The constitutive features of agency (no matter whether they are conceived as aims, motives, capacities, commitments, etc.) continue to operate even when the agent is assessing whether she is justified in her engagement in agency. One cannot put agency on hold while trying to determine whether agency is justified because this kind of practical reasoning is the exclusive job of intentional agency. This does not mean that agency falls outside the reach of reflection. But even reflection about agency is a manifestation of agency." (Ferrero, 2009, p. 309).

deliberately left vague by constructivists, so as not to provide a definitive or exhaustive list. Nonetheless, they are consistently conceived as constitutive constraints on deliberation and as prerequisites for participation in reasoning as *practical*.[730]

In what sense, then, is this not a foundational approach? Forst's principle to justification in particular appears to be one, as he himself seems to stress many times (Forst, 2012, p. 5).[731] What crucially distinguishes this form of constructivism from foundationalism, however, is that, first, as already noted, it rests on an epistemological rather than an ontological foundation, and secondly, constructivism does not exhibit a deductive structure: it does not aim to derive normative conclusions from first principles to pre-solve all the moral dilemmas we might face.[732] O'Neill effectively underlined how this is precisely one of the main Kantian root of constructivism, the rejection of any built-in "master plan" of reason and the claim that humans must themselves devise the guiding principles by which their faculties are assembled (O'Neill, 1989, pp. 18–19).[733] As she

---

[730] As even Scanlon emphasizes: "When we apply the method of reflective equilibrium to our moral beliefs, we assume hypothetically that our considered moral judgments represent conclusions about what we should do that are supported by an objective way of reasoning (determined and independent of persons) that all reasonable people have reason to regard as authoritative." (Scanlon, 2002, p. 147).

[731] "The "ultimate" foundation of constructivism cannot itself be constructed, but must prove itself as being appropriately reconstructed in an analysis of our normative world." (Forst, 2012, p. 5).

[732] Many non-constructivist authors also emphasize that practical reasoning does not operate as a closed deductive system, since there is no single, general decision procedure capable of determining the right action in every case. For example, this is a crucial point in Nagel's philosophical system, as highlighted in the quotation at the beginning of this chapter (Nagel, 1979, p. 135). For constructivists, however, the absence of a single ready-made deductive system is precisely why we need a constructive method to generate principles that agents can endorse as binding (Korsgaard, 1996b, pp. 35–36)

[733] "The claim that reason is nonalgorithmic sheds some light on Kant's anti-foundationalist textual strategy that postpones the discussion of reason until after the elaborate inventory of materials in the "Transcendental Doctrine of Elements". The "Doctrine of Elements" depicts cognitive capacities none of which can be deployed in abstraction from the others; yet it does not reach a determinate account of their integrated deployment. The story is familiar to all of Kant's readers: Human beings can receive a manifold of intuition, subject to the structural constraints of the forms of intuition. They can synthesize this manifold using empirical concepts in accordance with the categories of the understanding. However, they can deploy the categories in complete acts of judgment only if they adopt and follow certain Ideas of reason or maxims of judgment {CJ, V, 182) to organize this judging. In the end a practical principle must guide all complete acts of judgment. This is why at the end of the "Doctrine of Elements" we have only an "inventory" of disconnected "materials" and not a "construction". In making trial of the Copernican hypothesis we discovered that the complex capacities that we call sensibility and understanding are not usable without "Ideas of reason", that is, without the adoption of maxims to regulate the use of these capacities in thinking and acting. An account of human knowledge will be systematically indeterminate unless these maxims are identified and vindicated. Here we begin to see why Kant thinks that practical reason is fundamental to all reasoning, why there can be no complete rules for judging and why human reasoning is, as we might say, nonalgorithmic, down to the bottom. But we are still far from seeing which maxims for the integrated use of our more specific cognitive capacities have an

also puts it with lucid brevity: "[A] fuller account of practical reasoning must say *something* about processes of judgement and deliberation; but it cannot say everything. Ethical principles are not algorithms." (O'Neill, 1988, p. 14).

Rather than deriving substantive moral conclusions from first principles that would supposedly resolve all moral dilemmas in advance, constructivism instead provides a meta-normative criterion for identifying which considerations genuinely count as reasons. This standard can operate negatively: some putative reasons collapse on themselves, not only because they fail the test of universalizability, but when they are inconsistent and self-defeating as in Gibbard's well-known anorexia example (1990, pp. 165–166, 171–172).[734] At the same time, the constitutivist position gives us a positive criterion to detect reasons that flow from the very structure of agency itself and therefore cannot be discarded. The commitment involved in acting, for instance, is constitutive of our capacity to have reasons at all: treating it as optional would undermine the very possibility of action and defeat itself.

---

authority that justifies us calling them Ideas (or principles) of reason. The metaphor of the tribunal is used to make the point that we cannot expect the principles of reason to dominate and control all reasoned thought and action. Thought and action can at most be constrained, not fully determined by principles of reason. This does not mean that abstracted aspects of human reasoning are all nonalgorithmic; it is, after all, from certain abstracted modes of reasoning that we gain our understanding of what algorithms are. It follows only that any thinking or decision making that actually follows algorithms, whether or not it works within the limits of the supreme principle of reason, depends on some specific and perhaps groundless premises that permit a decisive selection of well-formed formulae and valid inferences. Algorithmic procedures taken in the abstract are incomplete, "private" uses of reason; formal calculi and models of rational choice presuppose rather than articulate what is most crucial for reasoning. Going back to the metaphors of construction, we might put it this way: The elements of human knowledge are not self-constructing; they must always be put together according to some plan or other. No master plan is inscribed in each one of us; rather we must devise a plan that assembles the various elements. This plan must not presuppose unavailable capacities to coordinate, such as a preestablished harmony between reasoners or between each reasoner and a transcendent reality." (O'Neill, 1989, pp. 18–19).

[734] This example was introduced by Allan Gibbard in order to show how agents can be practically incoherent within their own frame of reference. The anorexic person's reasoning can be seen as irrational because accepting "norms that prescribe death by starvation" for the pursuit of "a trim figure" is ultimately inconsistent (Gibbard, 1990, p. 171). For Gibbard, this kind of irrationality is not merely a disagreement with the anorexic person's goals or values; rather, it illustrates that some goals can undermine the practical rationality required for agency itself, much as logical contradictions undermine reasoning in formal logic. "Our so thinking consist in accepting a norm that says without qualification, "Do not starve yourself to death for the sake of a trim figure". That implies, among other things, such injunctions as "Do not starve yourself to death for the sake of a trim figure, even if you prefer to do so" and "Do not starve yourself to death for the sake of a trim figure, even if you think that doing so is rationally required" It is in senses like these that we regard the woman who starves for the sake of a trim figure to be acting irrationally." (Gibbard, 1990, p. 165).

If this seems to risk sliding back into formalism, it is crucial to restate one of constructivism's central insights: normativity does not stem from pre-given internal drives or external dictates, but rather emerges from a procedure that is at once constitutive and public, thereby securing the advantages of "both dimensions." (Forst, 2002, p. 260).[735] The constitutive principles of reason, though they frame and enable the very activity of reasoning, are not in themselves normative, because for constructivists all normative facts cannot be discovered or revealed, but only constructed through reason in response to particular practical contexts. Thus, even the constitutive principles acquire authority for us only through the reflective act of reasoning (subjectively), and they acquire validity for all only through deliberation with other reasonable agents (intersubjectively). I keep a promise not because of some prior metaphysical fact about the duty to obligation, but because, in reasoning about it, I recognize that (usually) no stronger reason overrides my keeping it, provided I am able to do so. In reaching this conclusion, I inevitably appeal, with more or less opacity, to one or more constitutive principles of reason (publicity, generality, universalizability).

This point regarding the authority of Reason is effectively illustrated by Raff Donelson through his thought experiment involving a modern oracle of Delphi (Donelson, 2017). Imagine a stone from outer space landing on Earth, that starts making factual predictions and issuing moral prescriptions. Initially met with skepticism, the stone—aptly named Delphi—quickly astonishes people by accurately revealing personal information and answering difficult questions across various fields, including history, science, and mathematics. As Delphi's vast knowledge leads to significant advancements, such as medical breakthroughs and space exploration, it gains widespread recognition as omniscient. However, doubts about Delphi's omniscience arise when it begins to make morally repugnant claims, endorsing practices like slavery and killing newborns for no reason (Donelson, 2017, p. 384).[736] This situation creates a divide among the populace:

[735] "The collective identity of an ethical community and the individual identity of members are therefore interwoven in both respects: answers to ethical questions must be compatible with both dimensions, but ultimately we must be able to assume responsibility for them individually. What matters here too is to overcome the alternative between atomism and social monism and to understand the phenomenon of intersubjectivity correctly: autonomous persons are intersubjectively constituted persons, and their autonomy is an autonomy "with" others, but they remain autonomous persons. Only as such can they understand themselves and be esteemed and recognized in communities as individuals." (Forst, 2002, p. 260).

[736] "Delphi makes some judgments that strike many as morally repugnant. "Enslaving other human beings is always morally permissible," says Delphi at one juncture. "Killing one's firstborn child is morally obligatory," it says

some question Delphi's omniscience in light of its controversial ethical assertions, while others find themselves swayed to accept its views over their own moral intuitions. Despite acknowledging Delphi's knowledge, no one feels compelled to act according to its disturbing moral principles. This illustrates that moral inquiry cannot be reduced to accepting a pre-established moral truth without understanding it. Delphi is not deliberating: it is merely issuing commands without explaining the reasons for them. And if we cannot identify a compelling reason to subscribe to an external moral truth, we cannot genuinely regard it as normative for ourselves.

This shows that if we cannot find a reason to subscribe to an external moral truth, we cannot make it genuinely normative for ourselves. Only through careful reasoning and reflection can we authentically endorse the normative authority of a moral principle (it is the first person singular that can reflectively confer such authority). Yet this is not merely a solipsistic task (O'Neill, 2000, p. 47),[737] for constructivism intrinsically incorporates the role of "others" in in the very constitution of moral norms and the authority they carry. The concept of reasons and values for constructivists, as we showed in section 2.2 are always relational in a certain sense, "something which implies the existence of another" since "normative claims are not the claims of metaphysical world of values upon us: they are claims we make on ourselves and each other (Korsgaard, 1993, p. 51).

The constructivist conception of moral authority is therefore able to respond to two distinct but related practical tasks. In the first-person singular, I must act on reasons that I can find normative and endorse as my own, grounded in my aim to constitute a unified, autonomous self. To properly act, I must see my actions as flowing from a law I give myself—one that reflects my practical identity. However, the question of why something is wrong in the first-person plural—why *we* ought not to do something—invokes a different dimension of normativity, one that presupposes the presence of fellow agents with equal claims to justification and autonomy. In this

---

some other time. Surely, some conjecture, this must be a prank or joke because the omniscient Delphi could not seriously believe such outrageous things. But the oracle isn't laughing. Others try to argue with Delphi about these alleged moral facts, but equanimous as it usually is, Delphi brusquely cites the assorted historical and sociological reasons that explain our commitments to other moral principles and then refuses to entertain more interlocutors." (Donelson, 2017, p. 384).

[737] "The activity of deliberation is not an exercise in introspection. Since Kant holds that agents are opaque to themselves – as to others – they may often be unsure whether a given principle is 'really' their maxim on a given occasion. But this is less important than may appear: even where knowledge fails, practice has a footing." (O'Neill, 2000, p. 47).

space, the authority is shared intersubjectively and mutually. To properly act, I must se my actions as flowing from a law that is not only self-endorsed but universalizable—one that could be willed by all. While these two domains are obviously interrelated,[738] to conflate them "is to conflate the normative source of moral obligations with the subjective authority they gain in the first person deliberative stance." (Bagnoli, 2021c, pp. 405–406).

Constructivism, understood in this light, avoids the pitfalls of both foundationalism and formalism. It rejects the queerness of foundationalism by denying that moral normativity rests on metaphysical truths or pre-given moral facts; instead, normativity arises from within the constitutive standpoint of agents who must justify their actions to themselves and to others. At the same time, constructivism resists the rigidity of strict formalism by grounding its procedure not in idealized rules alone, but in the practical conditions of mutual recognition among rational, vulnerable, and finite beings. Universality, then, is not an external truth to be discovered, but a constitutive and intersubjective construct that emerges from the ongoing practice of giving and asking for reasons among rational, finite, and vulnerable beings—beings who recognize one another as ends in themselves, and who are thereby also sources of justification or "normative justificatory beings" (Forst, 2024, p. 127).[739] Through its emphasis both on autonomy and constitutive consistency, and publicity and mutual recognition, constructivism is able to overcome the tensions between foundationalism and formalism without sacrificing normative force.[740]

---

[738] As reflective agents, we not only strive for coherence in our own agency, but we also care about how our reasons and actions appear to others. A maxim that can be endorsed from the standpoint of others—that is, one that is universalizable—often gains stronger authority for us in the first person, precisely because we understand ourselves as participants in a shared moral community.

[739] "This brief reflection on criteria of justification opens up the possibility of a stronger normative program. I call this a strong program because it operates with heavy normative vocabulary, which, however, is not grounded in a Platonic higher reality but – along Kantian lines – in a constructivist reflection on what it means to understand ourselves as normative justificatory beings. As such beings who use our faculty of reason to produce, test and accept or reject justifications, we face the task of constructing a common world of norms that are binding on and justifiable to all who are subjected to them. As agents of construction, we ought to regard ourselves as equal authorities of the norms that bind us all equally and, as such authorities, we respect each other's right to justification. The criteria of justification that hold between us as reasonable and responsible beings do not predetermine the content of our normative constructions, but they do determine the procedure of construction and thus also some content, given the respect we owe each other as constructive normative authorities. As such authorities, we are law-makers and law-takers at the same time, and thus the Kantian notion of dignity applies to us as subjects with rights and duties of justification." (Forst, 2024, p. 127).

[740] Constructivism is not the only approach that seeks to navigate a path between foundationalism and strict formalism. Pragmatism, for example, also rejects both extremes by emphasizing the practical consequences of our

5.3.2 Is Constructivist Universality Circular or Arbitrary? Transparency of Normativity

The many advantages of this constructivist conception of universality, achieved through the constitutivist move, may appear sufficient to demonstrate how the proposal distinguishes itself from the more extreme positions at either end of the spectrum. But what about the more targeted criticisms of circularity and arbitrariness which, as we saw in the previous chapter, are to some extent shared by otherwise opposing approaches like jusnaturalism and juspositivism? Can a constructivist conception of universality also withstand these challenges?

The standard objection is that the constructivist conception of universality is not, in fact, a genuine alternative: critics argue that constructivism must inevitably revert to some form of (arbitrary) realism (Shafer-Landau, 2003; Millgram, 2005; Enoch, 2006), or become circular in the selection of its norms (G. A. Cohen, 2003; Raz, 2003b; Timmons, 2003).[741] To translate these meta-ethical concerns into political terms and tie them to our broader inquiry, we might imagine a

moral beliefs and policies. Rather than grounding normativity in metaphysical truths or abstract procedures, pragmatists—like Richard Rorty—argue that moral and political principles should be justified by their capacity to address concrete problems, promote democratic deliberation, and enhance human flourishing. Justification, in this view, is contextual and outcome-oriented, not derived from ultimate foundations or rigid formal criteria. Similarly, Neo-Aristotelian Ethical Naturalism, as developed by philosophers such as Philippa Foot and Michael Thompson, avoids foundationalism by locating moral normativity not in a realm of transcendent values, but in the natural facts about human beings as living, rational, and socially embedded creatures. Moral evaluation, on this view, arises from the internal standards of human life forms and the capacities required for their flourishing. At the same time, it avoids strict formalism by rejecting purely procedural accounts of moral reasoning, instead grounding ethics in substantive conceptions of the good life and the virtues. Rather than seeking universalizability through abstract principles, this approach focuses on the cultivation of practical wisdom (*phronesis*) and context-sensitive moral judgment. While it should be clear how constructivism offers a distinctive framework for navigating the space between foundationalism and strict formalism, these various approaches can all be seen as complementary contributions to a broader philosophical project: the pursuit of a non-foundational and non-formalistic, yet normatively robust, philosophical account.

[741] Kantian constructivists were aware of these difficulties from the outset. Kant himself recognizes the apparent paradox that moral concepts, such as good and evil, are not determined prior to engaging in practical reasoning, but emerge only through it—a tension known as the *paradox of the method* (KpV, 5:62). However, O'Neill aptly emphasized how Kant does not rely solely on the familiar metaphor of a "tribunal of reason"—a metaphor that indeed risks making critique appear circular, with reason acting simultaneously as judge and accused—but also additional and crucial political metaphors to show how procedures of reasoning can possess genuine authority without presupposing their own validity. As O'Neill puts it: "Kant's account of the politics of reason does not lay everything on the metaphor of a tribunal. He uses other political metaphors to articulate further aspects of a critique of reason. Taken together they are to show why certain procedures for orienting thinking or acting may be said to have authority. In particular they help us to understand one of the puzzles with which we began: Is not any supposed critique of reason unavoidably circular, and so question-begging? (O'Neill, 1989, p. 20).

critique of constructivism in the context of human rights asserting that a "third way" is impossible: constructivism is either a form of jusnaturalism or juspositivism. As such, it is accused of being either an ungrounded foundationalism or a vicious formalism. If human rights are understood as positive rights, they become contingent upon human practices and cultural attitudes; this inevitably undermines their claim to universality and leads to circular reasoning in asserting their authority and internal coherence. If human rights are framed as moral rights, they appear arbitrary in their attempts to establish universality, relying on external source of authority that falls outside a formal procedural framework, ultimately resulting in an infinite regress of justification. Even when scholars like constructivist appeal to the concept of reason to get out of the impasse, the reasoning still remain either arbitrary, relying solely on individual perspectives without a broader foundation, or circular, when the vicious

Indeed, standard response to avoid the first critique is that what prevents the construction process from being arbitrary is considering reason as constitutive (Bagnoli, 2022a, p. 30), but many argue that this simply falls back into circularity (Millgram, 2022). This need not be the case, provided the arguments for the inescapability of agency hold. From a practical standpoint, if an agent begins to investigate whether their participation in agency is justified, he is not thereby committing a *petitio principii*, because, although the inquiry proceeds in accordance with the standards of agency, the agent does not yet know whether their "concern" for agency will ultimately prove justified. Simply, the very inescapability of agency implies that there is no alternative starting point from which such an investigation could be conducted.

While in certain more complex undertakings a vicious circularity can arise if one attempts to validate a rule of inference (e.g., *modus ponens*) by using that same rule, correctness-setting criteria for any enterprise (from playing chess to theoretical reasoning) cannot be "self-validated," precisely because their role is to define what counts as correct within that domain (Ferrero, 2009, pp. 324–325).[742] The same applies to agency: whether an action is valid within the "game" of

[742] "For some more complex enterprises, such as theoretical or practical reasoning, there is the possibility of a genuinely vicious circularity. One might reason theoretically about one's theoretical reasoning, or reason practically about one's practical reasoning. For instance, one might want to establish whether a rule of inference like modus ponens is theoretically valid. In doing so, one cannot rely on that very rule. This would be unacceptably circular. For there is an independently established criterion of theoretical correctness—conduciveness to the truth—that must be used to validate the rule of inference. To validate the rule of inference, one is to show, without relying on that very rule, that the rule meets the criterion of theoretical validity; that the rule is conducive to the truth. No similar question arises, however, for the criterion of theoretical validity, as opposed to a rule of inference. The criterion of theoretical

agency depends on agency's own constitutive standards, therefore trying to validate the fundamental criteria of practical reasoning in their own terms is not viciously circular but straightforwardly misunderstands their role. "The worry of circularity arises only because of the mistaken impression that the practical question is supposed to establish that the criteria are valid as the criteria that sets what counts as practically correct." (Ferrero, 2009, p. 326).[743]

Therefore, the seemingly inevitable circularity in the construction procedure is not a vicious one if properly conceived as constitutive of the practice. At the same time, the criteria that define a judgment as "considered" or "reasoned" emerge from within the deliberative process itself and does not therefore require external justification: they are built by the very activity of reasoning rather than imposed as "material starting points dictated by reality" " (Bagnoli, 2016, p. 1245).[744] This is the sense in which constructivism, by not being a deductive procedure, avoid the arbitrariness problem associated with foundationalism. The starting point is not chosen arbitrarily but exemplified by the transparency of the constitutive element of the procedure.

Here the distinctiveness of constructivism becomes evident in comparison with moral intuitionism. Intuitionism seeks to avoid arbitrariness by appealing to self-evident moral truths,

---

validity cannot be theoretically validated since its role is to determine what counts as being correct in theoretical reasoning." (Ferrero, 2009, pp. 324–325).

[743] "When one asks the practical question about agency—'Is there reason to be an agent?'—one relies on the criteria set by the nature of agency to answer this question given that this is a practical question. What one is asking is whether one should adopt these criteria in shaping one's conduct. One is not asking whether the criteria are valid in their status as setting what counts as correct in practical reasoning. This status is rather presupposed in raising the practical question. The status remains unaffected by the answer that one is going to give to the practical question. Like in the chess example above, the standards of agency continue to set what counts as practically correct whether or not one has any reason to adopt those standards. (Ferrero, 2009, p. 326).[743]

[744] "On the constructivist view, then, the effects of practical reasoning are quite remarkable and cannot be captured solely in terms of performing, bringing about something, realizing an end or producing a string of consequences. On this rich understanding, construction names a mode of reasoning genuinely alternative to classical models. First, it is principled but it does not derive normative conclusions from first principles; its structure is not deductive. Second, it starts from an abstract profile of practical subjects, concerned with others, but neither driven by altruistic concerns nor empowered with a distinctive faculty of moral insight. Third, there are constraints on the rational activities of construction, but they are structural of such activities rather than externally imposed on them. This means that there are criteria of correctness for constructing reasons, but such criteria are not material starting points dictated by reality. The domain of reasoning is built up by the activity of reasoning. This is the sense in which constructivism is not a form of realism. Fourth, the abstract profile of practical subjects includes beings with interests and concerns, responsive to the authority of objective reasons, even though these reasons are in some important respects independent of their pre-reflective desires and not pondered judgments. Constructivism does not represent the reflective agents as self-interested competitors trying to bargain and strike a deal. Its problem is to explain how objective reasons can be authoritative in the first person for subjects capable of practical concerns." (Bagnoli, 2016, p. 1245).

directly grasped through moral intuition, but it risks dogmatism, since claims of self-evidence are often disputed without further justification. Constructivism, by contrast, grounds morality in the constitutive features of agency, such that the starting point of moral reasoning is not arbitrarily chosen but arises from conditions no agent can coherently reject. In this way, arbitrariness is avoided not through intuition but through procedures that remain, in principle, accessible to all rational agents.

Although debates continue among constructivists regarding the role of intuition (O'Neill, 1989; Scanlon, 1998; Ronzoni & Valentini, 2013),[745] what unites them is the aim of explaining the authority of moral knowledge without appealing to an independent moral reality. In this respect, constructivism sets itself apart not only from intuitionism but also from projectivists or error theories, which regard evaluative discourse as systematically mistaken (Mackie, 1977; Blackburn, 1993). Constructivism's novelty lies precisely in explaining objective moral knowledge without presupposing an independent moral reality, but instead grounding it in reason's own activity. Practical reasoning does not uncover independent moral facts but instead yields a form of knowledge intelligible only through accounts that treat cognition as a self-conscious activity (Rawls, 2000, p. 138; Engstrom, 2009, pp. 7–8).

Among constructivist accounts, Stephen Engstrom arguably offers the most compelling direct argument for conceiving practical reason as a form of self-knowledge. In this model, the knowing subject and the object known—the agent—are originally united: practical knowledge is always the knowledge that individuals, as subjects who judge practically, have about how they themselves ought to act (Engstrom, 2013, pp. 145–146).[746]As practical knowers, individuals

[745] While O'Neill (1989) famously denied that intuitions play any justificatory role, Scanlon (1998) has argued forcefully that appeal to intuitions is unavoidable: justification proceeds by testing the coherence between theoretical assumptions and intuitive moral judgments, which function as initial data without themselves constituting independent truths. More recent developments in constructivism acknowledge the inescapable presence of intuitions in the phenomenology of moral experience and contend that the only way to preserve their significance without abandoning the public and intersubjective dimension of reason is precisely through a constructivist method that takes into account both intuitions and intersubjectivity. Ronzoni and Valentini (2013), for example, argue that this can be achieved by adopting the holistic method of reflective equilibrium, allowing for a limited and controlled use of intuitions.

[746] "Since the objects of theoretical knowledge exist independently of the knowledge, they must, if they are to be known, be somehow given, by affecting the senses; but an object I know through being affected by it I conceive, in the first instance, as other than myself. So theoretical knowledge is always of objects the cognizing subject first conceives as distinct from itself. Practical knowledge, in contrast, is a form of self-knowledge. As we just saw, practical judgment contains within its self-consciousness an understanding of itself as efficacious in respect of the

necessarily understand their own agency as the exercise of a capacity for rational self-knowledge—a condition that constitutes the form of the cognition their judgment aims to realize, and which serves as its internal and necessary standard of validity (Engstrom, 2013, pp. 146–147). The form of practical knowledge for Engstrom is therefore not an object of theoretical knowledge but resides exclusively in the original self-consciousness of the activity of practical judgment through which individuals constitute themselves as persons. The awareness of this form is internal to individuals' self-conception and to every practical judgment, yet it is distinct from and prior to any act of judgment since, and this is a central point in the constructivist proposal, our practical judgments begin from a conception of ourselves. To judge that one ought to φ implicitly involves claiming to know that one ought to φ, and thereby presupposes the capacity for such knowledge. Moral capacity is nothing other than the capacity for practical knowledge, and this insight resolves the long-standing debate over the constitution of fundamental moral capacity (Engstrom, 2013, p. 148).[747]

The constructivist appeal to transparency clarifies also why the rejection of arbitrariness does not come at the cost of circularity. The validity of the procedure derives not from arbitrary stipulation or covert reliance on independent moral facts, but from the transparent constitutive features of agency itself—features that no rational agent can coherently repudiate. And because constructivism is not a deductive enterprise, it does not attempt to demonstrate the validity of these standards from prior premises or from some neutral standpoint. Rather, they emerge from within the practice of reasoning in which we already inhabit. Constructivism is not about proving

---

action it represents. This understanding rests in a general understanding of the capacity for practical judgment as having an efficacy, an efficacy constituting the capacity to act in accordance with practical judgment. In the exercise of the capacity for practical judgment, then, the action represented as to be done is always understood as that same capacity's effi-cacy in operation. In practical knowledge, therefore, the knowing subject and the object known (the agent) are originally one. It follows that in practical knowledge the subjective and objective universal validity characteristic of all rational knowledge necessarily coincide. Practical knowledge is at once valid for all persons qua practically knowing subjects and valid of them all qua practically knowable objects, or agents. Practical knowledge is thus always knowledge persons, as practically judging subjects, have of how they themselves are to act, of what they themselves are to do." (Engstrom, 2013, pp. 145–146).

[747] "If the conception of the person figuring in ordinary practical thought is constituted, in the manner just suggested, by the form of practical knowledge, then a conception of the person is available that is both suitable for constructivism and congenial to objectivity in practical judgment. This form, it should be clear, is not to be understood along the lines of classical realism. It is not an object of rational perception or intuition nor anything that must be given from without in order to be known. Rather than being an object of theoretical knowledge, it lies solely in the original self-consciousness of the activity of practical judging through which persons constitute themselves as persons at all. Kant's idea that practical reason, as the basis of morality, has the character of autonomy (not heteronomy) underscores this point." (Engstrom, 2013, p. 148).

foundational rules but about making explicit what it already means to engage in reasoning or agency. To deliberate about action at all is already to occupy the standpoint of a reasoning agent who must justify her conduct in terms that others could, at least in principle, recognize. In this way, the authority of morality is not external or mysterious but remains intelligible and acceptable as transparent and accessible. Constructivism makes the source of moral authority visible by showing how it flows from conditions already implicit in agency, conditions we necessarily endorse as beings who reason and act in a shared world.

Thus, for constructivists, the procedure is transparent precisely because it is rooted in the unavoidable structure of agency itself. In this sense it is not opaque, since it does not appeal to anything external or mysterious; rather, it clarifies what is already inherent in our practical self-understanding. The charge of opacity is, however, not entirely baseless as long as we interpret it to means that we are not always fully transparent to ourselves, and we do not automatically recognize all the implications of our constitutive commitments as agents. There are limits to our self-knowledge, and our reasoning capacities can be clouded by bias, emotion, or ignorance. However, constructivism does not require perfect self-transparency; it requires only that the source of moral authority be in principle accessible through the procedures grounded in agency. The procedures remain intelligible because they flow from features of agency we cannot coherently reject, and, crucially, we always retain the capacity to reflect, deliberate, and improve our understanding of them. What might be opaque for a specific agent at any given moment is therefore not a fundamental mystery per se, but a challenge of practical self-reflection — the authority of moral principles remains open to recognition, clarification, and refinement through our ongoing exercise of reason.

If we recall again Delphi's thought experiment we see why even an oracle that delivered infallible truths would not automatically command our allegiance: without grasping and sharing the reasons for its prescriptions, we would not experience them as genuinely authoritative. Moral authority, in other words, is transparent only when it is both internal—rooted in our constitutive commitments as agents—and external—when these constitutive commitments are accessible within a shared space of reasoning. By anchoring authority in what is inescapably ours as reasoning agents, constructivism secures both intelligibility and normativity: the standards that guide us are not imposed from without, but arise from within the very activity of deliberation itself. Far from relying on arbitrary stipulation or succumbing to circularity, constructivism offers a notion of

universality rooted in the transparent features of agency that acknowledges the limits of self-knowledge, yet frames them not as insurmountable mysteries but as practical challenges—opportunities to reflect, deliberate, and refine our understanding. In doing so, constructivism illuminates a path that is both accessible and robust, capable of guiding action in a pluralistic world while preserving the normative force of universal claims.

5.4 The Universality of Human Rights in Constructivist Terms

I have argued that constructivism can retain the distinctive claim to universality in the human rights debate by offering a mode of justification that is neither arbitrary nor fully circular. It is not arbitrary insofar as it is anchored in the structure of practical rationality, which supplies normative standards that no agent can coherently repudiate while continuing to understand themselves as a rational actor. At the same time, it is not circular, since the norms generated by practical reason are not reducible to the contingent totality of existing beliefs, nor merely derived deductively from a fixed set of prior moral premises. Rather, they emerge from the reflective activity of agents who seek to justify their claims to one another under conditions of reciprocal recognition. It is precisely this justificatory structure that provides the normative basis for a constructivist account of human rights.

Accordingly, if at the end of Section I we arrived at a constructivist definition of human rights as reasonable demands that individuals address to one another on the basis of their mutually acknowledged moral standing as persons worthy of respect, we can now further ratify and refine this definition. The qualifier "reasonable" can in fact be understood in the universalist sense outlined above: a reasonable demand qualifies as a human right only insofar as it can be justified to all those subject to it on grounds that they, as rational agents who acknowledge one another as equals, cannot reasonably reject. Human rights are therefore universal demands: claims that all human beings can, in principle, recognize as valid for everyone. [748]

On this account, the universality of human rights does not rest on the postulation of a pre-social or metaphysical human essence. Instead, it arises from the reciprocal logic of moral

[748] Here the term "specificatory demands" has been shortened to simply "demands" for the sake of simplicity. It remains, however, that constructivists understand such demands as encompassing both the second-personal and specificatory dimensions emphasized by Darwall and Richardson.

recognition that structures practical reasoning itself and expresses the conditions under which agents can recognize one another as free and equal participants in a shared normative order. Once individuals acknowledge themselves and others as bearers of equal moral standing, they are committed, by a requirement of practical rational consistency, to extending the validity of such demands to all persons without exception. Any attempt to restrict these demands to particular groups, cultures, or communities would require justificatory reasons that undermine the very standpoint of equal recognition from which the claims are made.

A standard objection to this type of accounts concerns the status of those who deny the equal moral standing of others. If dignity is understood not as an objectively existing value but as a normative status constituted through practical recognition, it may seem that its binding force depends on widespread acknowledgment. What, then, are we to say when some individuals or social orders refuse such acknowledgment altogether? Human rights skeptics—ranging from racists to defenders of patriarchal hierarchies—often contest precisely the claim that all persons possess equal rational capacities, arguing instead that certain groups are less responsive to reasons or less capable of responsible agency. In short: if dignity depends on recognition, does its denial by some agents not undermine its universality?

This objection rests on a misunderstanding of the role recognition plays within the constructivist framework. Recognition is not a matter of empirical endorsement or subjective approval, but a normative commitment implicit in the very activity of practical justification. To engage in reason-giving at all—to assert claims, defend social arrangements, or justify exclusions—is already to presuppose a shared space of reasons in which others are addressed as potential respondents. From this standpoint, attempts to deny the moral standing of certain groups are not external challenges to constructivism, but internal failures of justification. Claims that some persons possess diminished rational capacities must themselves be justified by reasons that can be addressed to those affected. Yet such justifications invariably collapse into practical contradiction: they rely on the very capacities they deny, treating the excluded as beings to whom reasons are owed while simultaneously refusing to recognize them as entitled to reasoned justification. Any person who recognizes themselves as a moral agent must, by consistency, recognize others as such too. The skeptic's position thus undermines the normative standpoint from which it is articulated.

In this sense, constructivism need not claim that all individuals display identical level of moral competence. What matters is not the actual exercise of rational capacities in every case, but the possession of the normative status of being accountable to reasons. This status is not contingent upon intelligence, education, gender, race, or social role, but follows from the minimal capacity to be addressed as a responsible agent within a justificatory practice. To deny this status on the basis of contingent empirical traits is to substitute descriptive inequality for normative exclusion, a move that cannot be defended without abandoning the logic of public justifiability altogether. Therefore, the denial of recognition does not weaken the universality of human rights; on the contrary, such denial is precisely what calls human rights into operation. Human rights function not as reflections of existing moral consensus, but as critical standards against which social practices and justificatory claims can be assessed. They do not depend on actual mutual recognition, but on the unavoidable constitutive presuppositions of justification itself.

The fact that constructivism does not require actual consensus on the content of human rights, but only agreement on the standards of justification by which claims are assessed, also allows it to respond to a range of other familiar objections, including those grounded in cultural relativity, persistent moral disagreement, and ideological contestation. Disagreement, on this account, does not undermine universality; rather, it is one of the conditions under which the universalist aspiration of human rights becomes normatively meaningful. Persistent moral disagreement is compatible with constructivist universality so long as the parties involved continue to appeal to reasons that purport to be acceptable to all those affected in a reflective equilibrium of reasonable pluralism. Universality is not an empirical achievement to be measured by convergence of belief or practice, but a normative criterion that governs how claims ought to be justified and evaluated. It specifies the conditions under which power, authority, and social arrangements can be regarded as legitimate, even in contexts marked by deep pluralism and conflict. Human rights remain universal insofar as contested claims must be defended within a shared justificatory framework that presupposes the equal moral standing of all persons, regardless of whether such claims are in fact acknowledged or respected.

By addressing objections concerning denial of recognition, rational capacity, moral disagreement, and the influence of power, the constructivist reading clarifies both the scope and the limits of human rights. It neither collapses universality into mere consensus nor insulates it

behind metaphysical assumptions that remain inaccessible to public justification. Instead, it locates the normative force of human rights in the inescapable commitments of practical reason and reciprocal justification—commitments that bind agents precisely insofar as they seek to justify their actions and institutions to others.

CONCLUSIONS:

HUMAN RIGHTS AS UNIVERSAL DEMANDS

Why should the boundaries of states be viewed as presuppositions of justice rather than as institutions whose justice is to be assessed?

(O'Neill, 2000, p. 4).

Democracy, as I saw and see it, is a task: that of trying to take sensible – and possibly radical – steps to establish a practice of justification in which those subjected to norms, rules, and institutions have a real and effective possibility to become their authors. Is this realizable? I don't know. Is it worth trying? Yes, absolutely.

(Guillaume &
Forst, 2012, p. 107).

In this project, I have sought to explore how constructivist approaches can meaningfully contribute to the contemporary human rights debate. To that end, I have attempted to outline a coherent and unified interpretation of the various scholars who identify—explicitly or implicitly—as constructivists. This endeavor was full of significant practical challenges, given that divergences concerning the nature and function of human rights arise not only between opposing theoretical paradigms but also within the constructivist tradition itself. However, it is precisely because of these ongoing disagreements that I believe it is essential to develop and articulate a unified

framework in which certain key points are clearly established. Let us now briefly recapitulate what these points are.

In order to define the notion of human rights we started by defining the notion of rights in general, navigating the contemporary debates on the metaphysics of rights. While scholars often focus on either their nature, form, or function, typically in an exclusionary manner, where emphasizing one dimension obscures or conflates the others, in contrast we have shown how a coherent constructivist account can meaningfully address all three dimensions. We introduced and postponed the question regarding the nature of rights—whether they are primarily moral or legal—on the grounds that, for constructivists, this dichotomic framing is misguided. Morality and legality represent two distinct yet interconnected domains of normativity, and constructivists therefore reject the separation thesis, which posits a strict independence between the two, by emphasizing their normative congruence and overlap (Forst, 2002). Constructivists also accept that the primary logical form of a right is that of a claim, meaning that (claim-)rights necessarily correlate with (directed) duties. This does not imply that all duties have corresponding rights (O'Neill, 1996), but when such a correlation does exist, rights and duties are co-constituted—they emerge together through a relational structure—rather than being hierarchically ordered in favor of either the right or the duty (Richardson, 1990).

This relational aspect is inadequately captured by traditional theories, such as Will Theory or Interest Theory, which tend to overlook that the central function of a right is its role as a social demand for mutual accountability (Darwall, 2006; Forst, 2016; Richardson, 2018). This is precisely what Stephen Darwall captures with his notion of Demand. When we make a demand, we are not merely stating a claim or asserting a preference, but addressing another person as an equal; this implies a public framework in which agents recognize one another as having the authority to make claims and the responsibility to respond to them. When I say, for example, "I have a right to privacy" or "You have a duty to keep your promise", I am not simply describing a fact or expressing a desire. I am directly addressing another agent, demanding that they act in a way that can be justified to all those affected, thereby reaffirming our mutual accountability. Moreover, this accountability is not static but can be specified and extended by broader moral or political communities. It is precisely through this process of specification that one can distinguish between legitimate and illegitimate rights claims. The notion of an unjust right, then, is self-

defeating not because it contradicts the logical structure of a right (as a claim correlating to a duty), but because it contradicts the social function of rights—namely, to demand mutual accountability and justification within a shared moral or institutional order.

In the most general sense, then, constructivists conceive of rights as legitimate specificatory demands. In order to explain the specific kind of normativity that underpins this legitimacy, we delved more deeply into the possible nature of rights, their relationship to other normative concepts, and the broader connection between normativity, agency, and reasoning. Constructivists do not view normativity as arising from external or metaphysically objective moral facts. Instead, they ground it in autonomous reasoning—a Kantian idea in which individuals self-legislate through rational reflection (Korsgaard, 1996b). On this view, moral norms are binding not because they exist independently of us, but because they can be reasonably justified to oneself and to others. This justification operates on two levels: internally, through the constitutive norms of Reason (Bagnoli, 2021b; Korsgaard, 2019), and externally, through the public deliberation of norms within a moral community (Forst, 2024; Richardson, 2018). Constructivism does not begin with pre-given reasons or values; instead, it treats both as the outcomes of a dynamic and transformative process of practical reasoning. Reasons for action are not static or foundational—they are constructed conclusions of deliberation. Likewise, values are not intrinsic or mind-independent entities; they are constructed through reflection on their normative significance. In the same way, rights are not self-evident truths, but construction of reasons; not the starting point of the deliberation, but the end of it.

Therefore, for a constructivist, a right becomes normative and legitimate only when the demand it expresses can be reasonably justified to both the duty-holder and the right-holder, in light of the relevant material conditions and the constitutive principles of Reason. Rights should then be conceived as reasonable specificatory demands, where reasonableness is achieved through reflexive and shared reasoning processes. In this sense, the distinction between moral and legal rights, then, is not ontological but procedural—it lies in the type of justification each requires. Moral rights are grounded in general moral reasoning (i.e. the right to avoid suffering) while legal rights depend on contextual and institutional considerations (i.e. the right to vote at eighteen*)*. The constructivist framework emphasizes the open-ended and procedural task of specifying and

justifying rights, including identifying who holds legitimate claims and who is responsible for fulfilling the corresponding duties.

Who, then, are the beneficiaries of human rights, specifically? What distinguishes human rights beyond the truism that they are rights held by humans? To answer this, we investigated the notion of *dignity*, a concept frequently invoked in human rights documents and debates. Constructivists—contrary to many non-Kantian and even wrongly-Kantian interpretations—argue that dignity should not be understood as self-evident value (starting point of the deliberation) but as a status we understand we have reason to value in each individual as a rational and autonomous being who is entitled to respect. Constructivism, therefore, reverses the classical reading of dignity (Sensen, 2011, p. 213): individuals are not respected because they possess dignity as an intrinsic value; rather, they possess dignity because we recognize them as beings who must be respected. Recognizing this status means acknowledging that each person is capable of wronging and being wronged (Richardson, 2018, p. 84), giving and asking for (normative) reasons (Forst, 2012, p. 14), being accountable and taking responsibility for their actions (Bagnoli, 2007, p. 114), and being able of engaging in reciprocal relationships based on equality and mutual respect (Darwall, 2006, p. 245). In doing so, constructivists also reject the traditional opposition between instrumental (which ground human rights in their usefulness for promoting some other feature, like well-being) and non-instrumental justifications (which ground rights in the inherent worth of persons in an intrinsic way) but ground both their normative status (non-instrumental) and their functional role (instrumental) in a relational way.

Human rights can therefore be conceived as reasonable demands that individuals specify toward one another based on their mutually acknowledged moral standing (dignity) as persons worthy of respect. This understanding highlights the procedural and relational nature of rights as outcomes of practical reasoning among moral agents and it also clarifies how the notion of dignity can be conceived as a status attributable to every human being, based on their potential capacity to respond to the command of Reason and to be held responsible for reasons to act morally. But what about individuals such as young children or those with severe mental disabilities who may lack this capacity? Constructivists maintain that these individuals still possess rights—contrary to the views of some scholars (Wellman, 2011b). They do so either by appealing to the idea of mutual recognition based on potential humanity, even when that potential can no longer be realized (Forst,

1999; Darwall, 2022b)—or by grounding obligations in the reasoned norms of a moral community, which can recognize and advocate for the rights of those unable to articulate claims themselves (Bagnoli, 2007; Richardson, 2013). Can this approach be extended to other non-reciprocating agents, such as non-human animals? According to some constructivists, potentially yes (Richardson, 2018; Bagnoli, 2022a) while others remain more restrictive in this regard (O'Neill, 1997; Forst, 2024); nevertheless, even when extended beyond humans, a constructivist conception of person rights remains similar with the definition of human rights we just provided.

We then tried to further expand this constructivist definition of human rights through a clearer articulation of the additional features human rights are generally expected to possess, particularly their universality. To explore this idea, we began by examining the historical link often drawn between the universality of human rights and the tradition of natural rights. This led us to a critical reconstruction of the opposition between jusnaturalism and juspositivism, showing how this contrast has often been overstated and that many philosophers traditionally associated with one school often incorporate elements of the other. Kant was taken as a paradigmatic example: natural law theorists emphasize his moral foundations and the idea of universal law, while legal positivists focus on his account of law's coercive authority and his rejection of a right to revolution. These aspects are both present and they are not at odds, but are integrated through Kant's notion of practical reason, as scholars like Ripstein and Flikschuh have convincingly shown (Flikschuh, 2015; Ripstein, 2009).

We then extended the historical analysis into contemporary debates, showing that these too are still shaped by long-standing tensions within both traditions. Jusnaturalism, in particular, is challenged by the problem of queerness (how to justify moral properties as distinct from physical ones) and the problem of relativity (how to claim universality amid deep moral and cultural disagreement). The answers provided by jusnaturalists to these problems reveal their foundationalist shortcomings: in attempting to avoid arbitrariness by grounding law in supposedly universal moral truths, they inadvertently reintroduce arbitrariness through the uncritical selection of underlying metaphysical assumptions. For instance, John Finnis's list of seven basic goods—though presented as universally valid—offers little justification for those who do not accept them as self-evident (Finnis, 2011). The problem here, however, is not universalism per se, but jusnaturalism's tendency to ground it through foundationalist assumptions

Juspositivism faces related challenges, despite beginning from the opposite end of the spectrum, by rejecting the necessity of moral foundations for legal validity. Contemporary positivists maintain that they do not deny the objectivity or normativity of law—only its universality and its subordination to morality. For them, the validity of a law depends on its source, not on its moral content: legal institutions apply rules not because they would endorse them in a moral deliberative stage, but because their validity is already *authoritatively established* by social sources. Joseph Raz follows this line with his *Sources Thesis*, asserting that a law's validity depends solely on social facts (Raz, 1970). However, this position sits uneasily with his *Service Conception of Authority*, which justifies legal authority by claiming that it helps individuals better conform to moral reasons. Critics argue that the tension between these two positions remains unresolved and that by refusing to integrate moral deliberation—especially at the level of general principles—juspositivism risks collapsing into formalism: legitimizing itself without ever stepping beyond procedural criteria. In doing so, it weakens legal authority by rendering the notion of "unjust" or "immoral" law logically incoherent.

There is, however, an even more specific tension in legal theory: both jusnaturalism and juspositivism seem to struggle with arbitrariness and circularity. Jusnaturalism risks arbitrariness in asserting moral truths and circularity in justifying them, while juspositivism risks arbitrariness and circularity by grounding legal validity solely in procedural conformity. Therefore, both frameworks fall short of offering a publicly justifiable basis suitable for pluralistic democratic societies, where legitimacy requires laws to be understandable, contestable, and acceptable to free and equal citizens who hold diverse moral, cultural, and ethical perspectives. The final task of the project, then, was to articulate a constructivist conception of universality that avoids the metaphysical burdens of jusnaturalism and the empty procedure of juspositivism, while preserving both moral and political validity as well as practical feasibility.

Constructivist theories started precisely by recovering a Kantian conception of reason as neither metaphysical intuition nor instrumental calculation, but as the capacity for reciprocal justification under conditions of freedom and equality. Reason, in this view, is essentially public: it requires that principles be acceptable to all who are bound by them, not because they mirror antecedent truths, but because they can withstand the test of universalizability (G 4:461). In this sense, Rawls and Habermas, both pioneers of constructivism (albeit in different ways) sought to

develop these Kantian insights. Rawls conceived the original position to operationalize Kantian universality to explain how principles of justice can be acceptable to all under fair conditions. However, following *A Theory of Justice* and *Kantian Constructivism in Moral Theory*, Rawls moved toward a notion of objectivity, grounded in overlapping consensus, over that of universality, aiming for a framework less metaphysically demanding. Thus, in *Political Liberalism*, Kantian constructivism—which avoids metaphysical claims by securing objectivity through intersubjective reason—is seen by Rawls as a comprehensive doctrine, unsuitable for pluralistic societies; his new version of political constructivism, by contrast, confines itself to the political domain, grounding legitimacy in coherence and public agreement rather than metaphysical universality (Rawls, 1993, p. 116).

Habermas, by contrast, retains universality at the heart of his discourse ethics and criticizes Rawls' position as abstract and monological, failing to capture the dialogical nature of public reason. For Habermas, impartiality cannot be imposed by a hypothetical veil of ignorance; it must be continuously reaffirmed through actual public discourse. In order to do so, he systematically distinguishes morality, ethics, and political discourse, stressing their co-originality and the co-originality of rights and democratic participation. The "unforced force of the better argument" ensures that only reasoned justification—not power or coercion—determines outcomes (Habermas, 1998, p. 37). Within this framework, universality is formalized in the principle of universalization (U): a norm is valid only if it can be freely endorsed by all affected. (U) refines the weaker principle of discourse (D), grounding moral validity in intersubjective agreement rather than metaphysical truth: norms are valid only if they can secure consensus among participants in free and equal communication. Procedures are not valuable because they are self-enclosed or mechanically consistent, but because they embody the practice of co-authorship and mutual accountability. Law gains authority by being continuously tested, revised, and legitimated through the very practices of reason-giving that it structures.

The debate between Rawls and Habermas can be interpreted through the lens of the tension between jusnaturalism and juspositivism, and it is particularly striking how each philosopher accuses the other of embodying the very weaknesses associated with both traditions. Rawls criticizes Habermas for extending political theory into overly broad philosophical terrain, while Habermas maintains that political theory inevitably involves moral questions that cannot be

confined to private doctrines (Habermas, 2011, p. 112). At the same time, Habermas contends that Rawls's allegedly constructivist framework is in fact substantive rather than purely procedural, grounded in moral intuitions and marked by a residual foundationalism rooted in the jusnaturalist tradition. Rawls, conversely, accuses Habermas of advancing a form of procedural juspositivism by avoiding substantive moral commitments at the outset and relying too heavily on formal procedures to ground legitimacy (Rawls, 1995, p. 135); this, Rawls argues, risks compromising individual rights by implicitly prioritizing public autonomy and democratic will-formation over private autonomy, despite Habermas's explicit efforts to preserve both. In the end, both critiques carry weight: although Rawls and Habermas each seek to avoid metaphysical or naturalistic foundations for rights, their respective theories ultimately blend procedural and substantive elements, exposing them to the very risks of foundationalism and formalism they both strive to avoid (Finlayson, 2019, p. 165).

Therefore, if a constructivist conception of universality must provide a viable alternative to the limitations of both jusnaturalism and juspositivism (by avoiding the reintroduction of foundational assumptions that it initially set out to overcome, while also resisting the tendency to collapse into empty formalism) we need a model of universality that is both normatively robust and procedurally legitimate, capable of guiding action without appealing to fixed metaphysical truths or reducing morality to legal formality. This is the direction taken by contemporary constructivist scholars, who have sought to refine and expand the constructivist project in ways that preserve its procedural core while enriching its normative content.

Onora O'Neill, for example, by aptly distinguishing the test of universalizability (how one determines whether a principle can possess the property of universality) from heteronomous moral tests (such as the Golden Rule), elucidated how the charge that Kantian ethics is a form of empty formalism is misplaced: universality concerns what can be coherently willed for all, not what is or would be commonly shared (O'Neill, 1989, pp. 131–132). The normative grounding of such universality, however, does not depend on metaphysical oddities, but on the practical requirements of coherent agency—conditions that are neither purely logical nor transcendental, but rooted in the lived realities of embodied, situated agents. Her approach is negatively framed—rejecting principles such as coercion, deception, or violence that fail the test of universalizability and undermine agency—but it remains action-guiding and normatively robust, preserving the force of constructivism without returning to the metaphysical commitments of moral realism. Extending

this notion of universality to the political domain, O'Neill argues that justice must be cosmopolitan rather than confined to state borders. Against statist models, including Rawls's *Law of Peoples*, she insists that principles of justice must apply to all moral agents, regardless of nationality, and must be paired with clearly assigned obligations, including those of non-state actors (such as corporations and NGOs). Her conception of "reasonable universalism" thus leads to a form of institutional cosmopolitanism that is both principled and practicable, avoiding both utopian idealism and relativist retreat (O'Neill, 2000, p. 201).

Similarly, Rainer Forst's conception of reflexive universality advances a constructivist alternative that aims to overcome the binary between jusnaturalism and juspositivism. Forst insists on the need for a shared moral foundation, but one that emerges from a dialogical process of justification among free and equal individuals. His theory integrates the procedural demand for legitimacy with a substantive concern for inclusion, equality, and critique: by grounding normativity in reciprocity and generality, and by requiring that all affected can reasonably reject or accept norms, Forst's approach ensures that moral and political norms emerge both through recursive and dialogical reasoning individuals (Forst, 2024, pp. 199–200). This not only safeguards universality against arbitrariness and relativism, but also ties it to concrete human experiences, social contexts, and the lived realities of injustice and domination.

Therefore, O'Neill, Forst, and the broader contemporary constructivist tradition in general, demonstrate that constructivism can remain faithful to its procedural foundations while still providing substantive guidance for moral and political life. In doing so, they show that constructivist universality avoids the pitfalls of foundationalism and formalism, as well as arbitrariness and circularity. The constructivist notion of universality is neither externally imposed (as in foundationalism) nor vacuously procedural (as in formalism), but of practical significance due to the significance of its inescapability strictly connected with agency. In this way, constructivism also avoids arbitrariness (namely selecting principles without justification) and circularity (namely validating its procedures with their own standards). The charge of arbitrariness is deflected because the starting point—agency—is not optional but a transparent condition no rational agent can coherently reject, while the charge of circularity is mitigated by recognizing that the constitutive criteria of agency do not require independent validation: they define what counts as reasoning itself.

Constructivism, in short, secures universality in the human rights debate by offering a justification that is neither arbitrary (since it is anchored in the structure of rational action, which provides standards no agent can coherently repudiate) nor fully circular (since rational norms are not reducible to the contingent totality of beliefs or merely derived deductively from prior premises). At the same time, it resists the rigid polarities of a foundationalism externally imposed, through metaphysical oddities that exceed the reach of practical reason, and a merely internal formalism, lacking substantive content to ground the moral claims in meaningful and context-sensitive ways. The constructivist notion of universality surpasses both these false alternatives by integrating both an internal and an external dimension of normativity, which is self-recognized as inherently constitutive of human agency and intersubjectively constructed through processes of public deliberation.

Therefore, if the long-standing debate between jusnaturalism and juspositivism—both tied to the difficulties of foundationalism and formalism—still dominates much of the philosophy of human rights, then the way forward is to reframe the central issue. The real question is not whether human rights are "natural," but whether they are "universal", namely if they are reasonable specificatory demands that can claim universal validity in a pluralistic global context. Constructivism answers this by offering a more dynamic and flexible framework: universality is not treated as self-evident or absolute, but as the outcome of an ongoing rational discourse—one that respects the autonomy and perspectives of all participants, and seeks coherence and mutual justification rather than dogmatic certainty. This is the sense in which human rights can be conceptualized as "constructed" (instead of discovered) and yet retain their claim to universality. To assert that human rights are universal rights rather than natural rights means, from a constructivist perspective, not only to affirm their universality but specifically to claim that such universality is achieved in a procedural, not foundationalist, manner. This understanding allows us to refine our constructivist definition of human rights precisely as universal demands.

Does this mean that constructivism is ultimately an alternative to the first distinction highlighted at the start of this project (Maliks & Schaffer, 2017b) between the moral and political approaches to human rights? In a sense yes, but more accurately it shows why this rigid dichotomy is misplaced. Constructivism does not align with political approaches that claim the universality of human rights is grounded solely in commitments made by states through international law,

rather than in universal moral principles (Buchanan, 2015; Hannum, 2016).[749] Such an approach would leave us ill-equipped to challenge unjust laws or practices that violate basic human dignity. At the same time, constructivism avoids the metaphysically odd conceptions of universality attributed to some "moral" approaches (Finnis, 2011; Perry, 2000).[750] Instead, it provides a practical and reasoned framework for human rights—one that recognizes the critical role of positive implementation in ensuring their real-world relevance while maintaining normative rigor.

In this respect, the blending of moral and political considerations is a feature not exclusive to constructivism but also recognized by many non-constructivist thinkers traditionally categorized as either moral or political theorists. For instance, Tasioulas (2015)—widely considered a traditional moral theorist—emphasizes the strong moral foundation of human rights, independent of institutions, rooted in human dignity and justice; however, he also acknowledges the practical realities and cultural diversity that the discourse on human rights must navigate. Without abandoning universal principles, Tasioulas highlights the importance of pragmatic and institutional elements, arguing that human rights must be understood in relation to the political frameworks that implement them. For him, there is an essential interconnection between moral principles and the political conditions necessary to make human rights effective. Similarly, politically oriented scholars often propose an integrated view that combines moral justification with the practical role of human rights. For example, Beitz (2001) views human rights not as a consensus among all existing political doctrines or comprehensive worldviews but as principles

[749] Hannum specifically critiques the tendency to view human rights as universal moral solution to all social and political problems, leading to a proliferation of new rights that dilutes the core essence of those already recognized. He dramatically puts it in these words: "Human rights are on the verge of becoming a victim of their own success. Unless there is a conscious attempt to return to the principles of consensus and universality, the increasingly strident calls from European and other 'Western' human rights activists for adherence to the contemporary liberal European construct of society is likely to create a backlash in the rest of the world. This tendency is concurrently exacerbated by activists who see an expansive concept of 'rights' as the primary means to effect domestic social and political change, as well as by governments that wish to burden human rights with broader geopolitical initiatives. The result will be, ironically, to strengthen anti-human rights governments and others who have challenged the universal application of human rights by privileging cultural relativism over globally shared values." (Hannum, 2016, p. 413).

[750] Perry is skeptical about the ability of any secular morality to provide a coherent and robust foundation for the concept of human rights. As a result, he argues that the core premise of human rights—that every human being is "sacred" or "inviolable"—is inherently and unavoidably rooted in religious belief (Perry, 2000, pp. 15–32)

for international affairs that reasonable individuals, despite holding conflicting yet reasonable conceptions of the good life, could accept (Beitz, 2001, p. 276).

If we begin to view the distinction between moral and political approaches to human rights not as a rigid dichotomy but as a continuum of overlapping perspectives, constructivism—and other moderate positions—emerges as a middle ground, equidistant from either extreme. What unites many of these moderate approaches is a commitment to pluralist universalism: a conception of universality that is open-ended, reflexive, and grounded in human interaction rather than in any abstract or fixed essence (Sen, 2004; Donnelly, 2007; Corradetti, 2009).

Building on this, we arrive at the most abstract level of the debate, the opposition between absolutism and relativism, which can be seen as, respectively, the culminations of jusnaturalist and juspositivist approaches (or, alternatively, of the moral and political approaches) when considered in their most antinomic forms.[751] Absolutism insists that human rights apply everywhere, regardless of circumstances; relativism, by contrast, concludes that genuine universality is impossible because every culture is grounded in its own distinct framework. A constructivist conception of universality allows us thus to explain why a law can be unjust (if it violates constitutive principles of human reason) without falling into forms of metaphysical absolutism (such principles are not mind-independent and are not absolute, but they are also not arbitrary). Similarly, it is possible to acknowledge the rich spectrum of cultural practices without falling into forms of cultural relativism, recognizing that underlying such practices are those constitutive principles of reason, as a cross-cultural and cross-generational matrix.

This combination of flexibility and normativity makes the constructivism approach well-suited for human rights in a globalized world. In an international context, where diverse political, legal, and ethical systems must interact, a rigid appeal to natural law may fail to resonate universally. Constructivism, by contrast, framing human rights as a practical effort in which all

[751] Whether one agrees or not with the labels used to categorize these currents, the aim here is to identify the elements that define the methodological divergences highlighted by Schaffer and Maliks in the current debate on human rights have. In this project, I have tried to highlight how these oppositions can be understood at different levels: the more specific concerns that these currents seem to share, such as the risk of arbitrariness or circularity; the more precise theoretical problems associated with foundationalism and formalism; and the broader risks represented by absolutism and relativism. This progression allows us to trace the logical and methodological conceptual tensions that underpin the debate, moving from the academic dispute to their more concrete implications in the public space.

individuals or agents are given a voice and engage in reasoned deliberation, recognizes the complexity of the global moral landscape while still maintaining a commitment to the idea that certain norms—such as human rights—can hold universally. By anchoring universality in a truly public and transformative Reason rather than in immutable facts, this approach is particularly effective in addressing some of the classic challenges posed by relativism against the globalization and the human rights project. Laws or practices that fundamentally undermine the dignity or rational autonomy of individuals—such as laws enforcing discrimination or subjugation—can be critiqued not because they violate external objective truths, but because they fail to align with the constitutive principles of human reason that we recognize guiding fair and just action across all societies.

In doing so, this conception of universality not only resists the charge of cultural relativism but also provides a coherent response to cultural imperialism (Mutua, 2000), demonstrating that it is possible to respect cultural specificity while upholding universally valid principles grounded in rational deliberation. The American Anthropological Association for example states that it is impossible to find a lowest common denominator, a collective heritage, between different cultures and thus the adjective "universal" is at best misleading and at worst dangerous, since it risks turning one culture's values into a standard imposed on all others. In their words: "Standards and values are relative to the culture from which they derive so that any attempt to formulate postulates that grow out of the beliefs or moral codes of one culture must to that extent detract from the applicability of any Declaration of Human Rights to mankind as a whole" (Association of American Anthropologist 1947, second proposition).

This statement seems to overlook the important distinction between norms that are merely created and those that are constitutive of human action. Without this critical distinction, every norm or right is rendered inherently cultural and contingent, which leads to the unpractical conclusion that universality is unattainable, leaving us with no conceptual tools to properly critique systemic injustices such as slavery, apartheid, or corruption, which risks thus being normalized within specific traditions or power structures. Such a misunderstanding narrows the horizon of human rights theory, preventing the development of a framework that both respects cultural diversity and affirms principles of genuinely universal scope. Constructivism, by contrast, clarifies

this distinction and thereby makes possible a vision in which cultural pluralism can flourish without obstructing the pursuit of global justice and universal human rights.

Recall some of the examples we explored in the introduction, such as famine or female circumcision. We can now see more clearly how a constructivist conception of human rights is equipped to address such questions by reasoning from shared human experiences toward rational moral thresholds. To deepen our understanding, let us introduce, before concluding, another complex and widely debated issue: child labor. Across cultures, children's participation in work can take vastly different forms. In some societies, it might be regarded as a valuable part of socialization, cultural transmission, or economic support;[752] but in order to distinguish legitimate traditions of participation of children to work from forms of exploitative labor that deprives children of education, safety, and development,[753] the only way to deliberate appropriately is by reasoning recursively and publicly on the specific question at hand. Exposing children to hazardous work, long hours, or abuse clearly violates the rational moral threshold: no reasonable justification can defend sacrificing a child's future for profit or tradition. Cultural specific values must be respected only as long as they don't undermine the dignity of children. International conventions (like the UN Convention on the Rights of the Child and ILO agreements) mandate laws prohibiting child labor and ensuring access to education precisely on these grounds, and this is how Human rights translate moral concern into enforceable legal obligations. When child labor persists, local and international activists, as well as NGOs and Intergovernmental Organizations,

[752] In nomadic or pastoralist societies, like Mongolia or some parts of Kenya, children often assist in herding animals as part of family life and cultural learning. Similarly, rural areas in France allow minors aged 14+ to help with seasonal farm work (especially *vendanges)* during school holidays under strict regulations and parent permission.

[753] In the Democratic Republic of Congo (DRC), which supplies over 70% of the world's cobalt (used in batteries for smartphones and electric vehicles), thousands of children—some as young as 7—work in hazardous, unregulated artisanal mines, where they dig deep tunnels, carry heavy loads, and are regularly exposed to toxic cobalt dust, all without protective equipment or access to education. Even in the United States, troubling instances of unjust child labor persist. As Human Rights Watch has meticulously documented (https://www.hrw.org/news/2023/05/04/us-should-end-child-labor-agriculture) agricultural labor laws in the U.S. contain unique exemptions that allow children as young as 12 years old—and sometimes even younger on small farms—to work unlimited hours outside of school. This legal loophole disproportionately affects undocumented migrant children, many of whom are employed under grueling and hazardous conditions: harvesting crops in extreme heat, handling dangerous pesticides, and facing a high risk of illness and injury, all without the protections typically afforded in other sectors.

are justified in invoking human rights to hold governments and corporations accountable, push for reforms, and demand fair labor practices globally.

Here we see how human rights reclaim both their moral and political dimensions —not as two distinct levels, but as mutually reinforcing elements—capable of enabling genuine systemic change. In the constructivist framework, human rights must first be understood as a regulative idea in the Kantian sense: not as fixed or self-evident truths, but as guiding principles at which we arrive through the open-ended practical task of specifying and justifying rights, including what duties they entail and who the legitimate duty-bearers are. In this way, human rights provide a practical normative horizon that not only orients our moral aspirations but also shapes political practice, offering a standard against which institutions and social arrangements can be critically assessed, equipping activists, NGOs, and citizens with normative tools to mobilize and hold power to account, while at the same time guiding states in the design and reform of laws and policies.

Nevertheless, one of the primary objections that are usually raised against a constructivist account is that it represents a revisionist approach, potentially undermining or impoverishing this empowering capacity and the geopolitical infrastructure that has been meticulously constructed over the course of decades. Critics argue that this infrastructure is not merely a theoretical construct but the product of specific historical contingencies but rather the result of specific historical contingencies—particularly the post-Second World War order—that have shaped the global system we recognize today. These contingencies encompass seminal events such as the establishment of the United Nations, the formulation of the Universal Declaration of Human Rights, and are regarded as hard-won achievements, forged in response to unprecedented global crises, which are unlikely to resurface under similar circumstances. The concern, then, is that adopting a constructivist approach could undermine this cohesive vision, historically guiding global efforts to address shared challenges, and risk producing a more fragmented or less effective

This objection, however, seems to miss the mark when we seriously consider the mounting evidence of the gradual decline in the significance of how the concept of human rights is utilized in global discourse. Once a powerful and unifying idea, human rights have increasingly become a contested and diluted framework, subject to competing interpretations and instrumentalization by various actors on the global stage. This erosion of coherence and impact calls for a critical

reassessment of the foundational assumptions and practices surrounding human rights to reinvigorate their relevance and effectiveness. A critical philosophical approach should expose these contradictions and allow them to "explode" (Adorno, 1973).

Yet, we cannot stop at this *pars destruens* (the deconstructive phase); to move forward, we must also engage with the *pars construens* (the constructive phase), namely building a solid framework that can sustain and advance the notion of human rights in the face of both longstanding and emerging critiques. By doing so, human rights can reclaim their dual function as both a moral compass and a political lever, addressing systemic injustices while empowering civic and institutional intervention. I leave it to future work to explore in detail how this framework can be further operationalized and, in particular, how Kant's vision of a "republic of republics" should be interpreted—as a fully institutionalized transnational entity, a coalition of sovereign states, or some combination of these possibilities.

REFERENCES

Adorno, T. W. (1973). *Negative Dialectics* (E. B. Ashton, Trans.). Suhrkamp.

Allan, T. R. S. (2005). *Constitutional justice: A liberal theory of the rule of law*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199267880.001.0001

Allison, H. E. (1986). Morality and Freedom: Kant's Reciprocity Thesis. *The Philosophical Review*, *95*(3), 393–425. https://doi.org/10.2307/2185466

Alston, W. P. (1976). Two Types of Foundationalism. *The Journal of Philosophy*, *73*(7), 165. https://doi.org/10.2307/2025920

Alvarez, M., & Way, J. (2024). Reasons for Action: Justification, Motivation, Explanation. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2024). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2024/entries/reasons-just-vs-expl/

Anderson, E. (1995). *Value in ethics and economics* (2. print). Harvard Univ. Press.

Anscombe, G. E. M. (1956). Intention. *Proceedings of the Aristotelian Society*, *57*, 321–332.

Anscombe, G. E. M. (1958). Modern Moral Philosophy. *Philosophy*, *33*(124), 1–19.

Applbaum, A. I. (2001). *Ethics for Adversaries: The Morality of Roles in Public and Professional Life*. Princeton University Press.

Arendt, H. (1959). *The Human Condition*. Anchor Books.

Association of American Anthropologist. (1947). Statement on Human Rights. *American Anthropologist*, *49*(4), 539–543. https://doi.org/10.1525/aa.1947.49.4.02a00020

Austin, J. (1832). *The Province of Jurisprudence Determined* (W. E. Rumble, Ed.). Cambridge University Press.

Baggini, J. (2018). *How the world thinks: A global history of philosophy*. Granta.

Bagnoli, C. (2002). Moral Constructivism: A Phenomenological Argument. *Topoi*, *21*(1), 125–138. https://doi.org/10.1023/A:1014805104487

Bagnoli, C. (2005). Breaking Ties: The Significance of Choice in Symmetrical Moral Dilemmas. *Dialectica*, *60*(2), 157–169.

Bagnoli, C. (2006). Humanitarian intervention as a perfect duty: A Kantian argument. *NOMOS*, *47*, 117.

Bagnoli, C. (2007). Respect and Membership in the Moral Community. *Ethical Theory and Moral Practice*, *10*(2), 113–128. https://doi.org/10.1007/s10677-006-9053-3

Bagnoli, C. (Ed.). (2011). *Morality and the emotions*. Oxford University Press.

Bagnoli, C. (Ed.). (2013a). *Constructivism in Ethics*. Cambridge University Press. https://doi.org/10.1017/CBO9781139094221

Bagnoli, C. (2013b). Reason and Ethics. In M. C. Amoretti & N. Vassallo (Eds.), *Reason and Rationality* (pp. 111–128). De Gruyter. https://doi.org/10.1515/9783110325867.111

Bagnoli, C. (2014). Morality as Compromise vs. Morality as a Constraint. *International Journal of Applied Philosophy*, *28*(1), 159–169. https://doi.org/10.5840/ijap20145621

Bagnoli, C. (2015). Moral Objectivity: A Kantian Illusion? *The Journal of Value Inquiry*, *49*(1–2), 31–45. https://doi.org/10.1007/s10790-014-9448-7

Bagnoli, C. (2016). Kantian Constructivism and the Moral Problem. *Philosophia*, *44*(4), 1229–1246. https://doi.org/10.1007/s11406-016-9745-4

Bagnoli, C. (2017). Kant in Metaethics: The Paradox of Autonomy, Solved by Publicity. In M. C. Altman (Ed.), *The Palgrave Kant Handbook* (pp. 355–377). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-54656-2_16

Bagnoli, C. (2018a). Defeaters and practical knowledge. *Synthese*, *195*(7), 2855–2875. https://doi.org/10.1007/s11229-016-1095-z

Bagnoli, C. (2018b). Values. In G. Bongiovanni, G. Postema, A. Rotolo, G. Sartor, C. Valentini, & D. Walton (Eds.), *Handbook of Legal Reasoning and Argumentation* (pp. 139–171). Springer Netherlands. https://doi.org/10.1007/978-90-481-9452-0_6

Bagnoli, C. (2019). Authority as a contingency plan. *Philosophical Explorations*, *22*(2), 130–145. https://doi.org/10.1080/13869795.2019.1599052

Bagnoli, C. (2021a). The objective stance and the boundary problem. *European Journal of Philosophy*, *29*(3), 646–663. https://doi.org/10.1111/ejop.12678

Bagnoli, C. (2021b). The Practical Significance of Kant's Categorical Imperative. In C. Bagnoli, *Oxford Studies in Normative Ethics Volume 11* (pp. 177–198). Oxford University Press. https://doi.org/10.1093/oso/9780192856913.003.0010

Bagnoli, C. (2021c). The Sources and Stances of Moral Normativity. *Etica e Politica*, *23*(2), 397–414. https://doi.org/10.13137/1825-5167/32392

Bagnoli, C. (2022a). *Ethical Constructivism*. Cambridge University Press. https://doi.org/10.1017/9781108588188

Bagnoli, C. (2022b). Hard Times: Self-Governance, Freedom to Change, and Normative Adjustment. In *Time in Action*. Routledge.

Bagnoli, C. (2022c). The Form of Practical Reasoning. In F. Ademollo, F. Amerini, & V. De Risi (Eds.), *Thinking and Calculating* (pp. 433–450). Springer International Publishing. 10.1007/978-3-030-97303-2_22

Bagnoli, C. (2024). Constructivism in Metaethics. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2024). Metaphysics Research Lab,

Stanford University. https://plato.stanford.edu/archives/win2024/entries/constructivism-metaethics/

Bagnoli, C. (2025). Evil and the Powers of Shared Agency. *Philosophical Topics*.

Baker, H. C. (1947). *The Dignity of Man. Studies in the Persistence of an Idea*. Harvard University Press.

Baron, M. W. (2008). Kantian Ethics. In M. W. Baron, P. Pettit, & M. A. Slote (Eds.), *Three methods of ethics: A debate* (11. Auflage). Blackwell.

Becker, L. C. (1980). Three types of rights. *Georgia Law Review*, *13*.

Beitz, C. R. (2000). Rawls's Law of Peoples. *Ethics*, *110*(4), 669–696. https://doi.org/10.1086/233369

Beitz, C. R. (2001). Human Rights as a Common Concern. *The American Political Science Review*, *95*(2), 269–282.

Beitz, C. R. (2004). Human rights and the law of peoples. In *The Ethics of Assistance: Morality and the Distant Needy* (pp. 193–214). Cambridge University Press. https://doi.org/10.1017/CBO9780511817663.011

Beitz, C. R. (2009). *The Idea of Human Rights*. Oxford University Press.

Benhabib, S. (1986). *Critique, norm, and utopia: A study of the foundations of critical theory*. Columbia Univ. Press.

Benhabib, S. (1992). *Situating the Self: Gender, Community and Postmodernism in Contemporary Ethics*. Routledge. https://doi.org/10.4324/9781003059516

Benhabib, S. (2015). [Review of *The Right to Justification*, by R. Forst]. *Political Theory*, *43*(6), 777–837.

Bentham, J. (1789). *The Collected Works of Jeremy Bentham: An Introduction to the Principles of Morals and Legislation* (J. H. Burns & H. L. A. Hart, Eds.). Oxford University Press. https://doi.org/10.1093/oseo/instance.00077240

Besch, T. M. (2020). Forst on Reciprocity of Reasons: A Critique. *The Southern Journal of Philosophy*, *58*(3), 357–382. https://doi.org/10.1111/sjp.12386

Biasetti, P. (2015). Infinite Regress and Hohfeld: A Comment on Hillel Steiner's "Directed Duties and Inalienable Rights." *Ethics*, *126*(1), 139–152. https://doi.org/10.1086/682189

Blackburn, S. (1993). *Essays in quasi-realism*. Oxford Univ. Press.

Bobbio, N. (2014). *L'età dei diritti*. EINAUDI.

Bourdieu, P. (1990). *The logic of practice*. Stanford Univ. Press.

Boyd, R. (1980). Scientific Realism and Naturalistic Epistemology. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *1980*, 613–662.

Boyd, R. (1988). How to Be a Moral Realist? In G. Sayre-McCord (Ed.), *Essays on moral realism* (pp. 181–228). Cornell University Press.

Boyd, R., & Richerson, P. J. (1988). *Culture and the Evolutionary Process*. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/C/bo5970597.html

Boyle, J. (1987). Thomas Hobbes and the Invented Tradition of Positivism: Reflections on Language, Power, and Essentialism. *University of Pennsylvania Law Review*, *135*(2), 383–426. https://doi.org/10.2307/3312102

Brink, D. O. (1989). *Moral Realism and the Foundations of Ethics* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511624612

Brink, D. O. (1992). A Puzzle About the Rational Authority of Morality. *Philosophical Perspectives*, *6*, 1–26. https://doi.org/10.2307/2214236

Broome, J. (2001). Normative Practical Reasoning. *Aristotelian Society Supplementary Volume*, *75*(1), 175–193. https://doi.org/10.1111/1467-8349.00085

Broome, J. (2013). *Rationality through reasoning*. Wiley Blackwell.

Brudner, A. (2011). Private Law and Kantian Right. *University of Toronto Law Journal*, *61*(2), 279–311. https://doi.org/10.3138/utlj.61.2.279

Buchanan, A. (2010). The Egalitarianism of Human Rights. *Ethics*, *120*(4), 679–710. https://doi.org/10.1086/653433

Buchanan, A. (2011). *Beyond humanity ? The ethics of biomedical enhancement*. Oxford university press.

Buchanan, A. (2015). Why International Legal Human Rights? In R. Cruft, S. M. Liao, & M. Renzo (Eds.), *Philosophical Foundations of Human Rights* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199688623.003.0014

Buchanan, A. (2017). *The Heart of Human Rights*. Oxford University Press.

Byrd, B. S. (2010). Intelligible Possession of Objects of Choice. In L. Denis (Ed.), *Kant's Metaphysics of Morals: A Critical Guide*. Cambridge University Press.

Callanan, J. J., & Allais, L. (Eds.). (2020). *Kant and Animals*. Oxford University Press.

Callicott, J. B. (1989). *In defense of the land ethic: Essays in environmental philosophy*. State University of New York Press.

Caranti, L. (2011). Kant's theory of human rights. In *Handbook of Human Rights*. Routledge.

Caranti, L. (2018). Reply to my critics. *Revista de Estudios Kantianos*, *3*(1), 110. https://doi.org/10.7203/REK.3.1.12335

Caranti, L. (2019a). *Kant's Political Legacy: Human Rights, Peace, Progress*. University of Wales Press. https://press.uchicago.edu/ucp/books/book/distributed/K/bo26510890.html

Caranti, L. (2019b). The dignity approach to human rights and the impaired autonomy objection. *Human Affairs*, *29*(3), Article 3. https://doi.org/10.1515/humaff-2019-0023

Caranti, L. (2023). On the Distinction between Humanity and Personality in Kant. In *The Kantian Subject*. Routledge India.

Caranti, L. (Forthcoming). *Kant, The Realist*. *Manuscript presented during Rainer's Forst Kolloqium the 4th of February 2024*.

Cohen, G. A. (1996). Reason, humanity, and the moral law. In C. M. Korsgaard (Ed.), *The Sources of Normativity* (pp. 167–188). Cambridge University Press. https://doi.org/10.1017/CBO9780511554476.007

Cohen, G. A. (2003). Facts and Principles. *Philosophy & Public Affairs*, *31*(3), 211–245. https://doi.org/10.1111/j.1088-4963.2003.00211.x

Cohen, J. (2004). Minimalism About Human Rights: The Most We Can Hope For? *Journal of Political Philosophy*, *12*(2), Article 2. https://doi.org/10.1111/j.1467-9760.2004.00197.x

Cohen, J. (2006). Is There a Human Right to Democracy? In C. Sypnowich (Ed.), *The Egalitarian Conscience: Essays in Honour of G. A. Cohen*. Oxford University Press.

Coleman, J. L. (2003). *The practice of principle: In defence of a pragmatist approach to legal theory*. Oxford Univ. Press.

Coleman, J. L. (2009). Beyond Inclusive Legal Positivism. *Ratio Juris*, *22*(3), 359–394. https://doi.org/10.1111/j.1467-9337.2009.00430.x

Copp, D., & Morton, J. (2022). Normativity in Metaethics. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2022). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2022/entries/normativity-metaethics/

Corradetti, C. (2009). *Relativism and human rights: A theory of pluralistic universalism*. Springer.

Corradetti, C. (ed.). (2011). *Philosophical dimensions of human rights: Some contemporary views*. Springer Science & Business Media.

Cruft, R. (2004). Rights: Beyond Interest Theory and Will Theory? *Law and Philosophy*, *23*(4), 347–397. https://doi.org/10.1023/b:laph.0000015417.05737.0e

Cruft, R. (2006). Why aren't duties rights? *The Philosophical Quarterly*, *56*(223), 175–192. https://doi.org/10.1111/j.1467-9213.2006.00436.x

Cruft, R. (2019). *Human rights, ownership, and the individual* (First edition). Oxford University Press.

Cruft, R., Liao, S. M., & Renzo, M. (Eds.). (2015). *Philosophical Foundations of Human Rights*. Oxford University Press.

Dancy, J. (2002). *Practical reality*. Oxford University Press.

Dancy, J. (2004). *Ethics without principles*. Clarendon Press.

Dancy, J. (2005). Should We Pass the Buck? In T. Rønnow-Rasmussen & M. J. Zimmerman (Eds.), *Recent Work on Intrinsic Value* (pp. 33–44). Springer Netherlands. https://doi.org/10.1007/1-4020-3846-1_5

Darwall, S. (1983). *Impartial reason*. Cornell University Press.

Darwall, S. (1990). Autonomist Internalism and the Justification of Morals. *Noûs*, *24*(2), 257–267. https://doi.org/10.2307/2215527

Darwall, S. (2006). *The second-person standpoint: Morality, respect, and accountability*. Harvard University press.

Darwall, S. (2007). Reply to Korsgaard, Wallace, and Watson. *Ethics*, *118*(1), 52–69. https://doi.org/10.1086/522018

Darwall, S. (2012). Bipolar Obligation. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics Volume 7* (pp. 333–358). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199653492.003.0011

Darwall, S. (2013a). Grotius at the Creation of Modern Moral Philosophy. In S. Darwall (Ed.), *Honor, History, and Relationship: Essays in Second-Personal Ethics II* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199662609.003.0009

Darwall, S. (2013b). *Honor, History, and Relationship: Essays in Second-Personal Ethics II* (1st edition). Oxford University Press.

Darwall, S. (2013c). *Morality, Authority, and Law: Essays in Second-Personal Ethics I*. Oxford University Press.

Darwall, S. (2022a). Animal Value and Right. In T. Schapiro, K. Ebels-Duggan, & S. Street (Eds.), *Normativity and Agency: Themes from the Philosophy of Christine M. Korsgaard* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780198843726.003.0010

Darwall, S. (2022b). Williams, Kant, and Morality's "Peculiarity." In A. Szigeti & M. Talbert (Eds.), *Morality and Agency* (1st ed., pp. 87-C4.P106). Oxford University PressNew York. https://doi.org/10.1093/oso/9780197626566.003.0005

Darwall, S., Gibbard, A., & Railton, P. (1992). Toward Fin de siècle Ethics: Some Trends. *The Philosophical Review*, *101*(1), 115–189. https://doi.org/10.2307/2185045

Dewey, J. (1925). *Experience and Nature*. Open Court. https://www.abebooks.com/first-edition/Experience-Nature-DEWEY-John-Open-Court/30611768081/bd

Diels, H. (2019). *Die Fragmente der Vorsokratiker: Griechisch und deutsch* (K. Walter, Ed.). Cambridge University Press.

Donelson, R. (2017). Ethical Pragmatism. *Metaphilosophy*, *48*(4), 383–403. https://doi.org/10.1111/meta.12253

Donnelly, J. (2002). *Universal Human Rights in Theory and Practice*. Cornell University Press.

Donnelly, J. (2007). The Relative Universality of Human Rights. *Human Rights Quarterly*, *29*(2), 281–306.

Doris, J. M., & Plakias, A. (2008). How to argue about disagreement: Evaluative diversity and moral realism. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol 2: The cognitive science of morality: Intuition and diversity* (pp. 303–331). Boston Review.

Dreier, J. (2004). Decision Theory and Morality. In A. R. Mele & P. Rawling (Eds.), *The Oxford Handbook of Rationality* (pp. 156–179). Oxford University Press. https://doi.org/10.1093/0195145399.003.0009

Driver, J. (2007). *Ethics: The fundamentals*. Blackwell.

Duchesneau, F., LaFrance, G., & Piche, C. (Eds.). (2000). *Kant actuel: Hommage à Pierre Laberge*. Bellarmin.

Dworkin, R. (1967). The Model of Rules. *The University of Chicago Law Review*, *35*(1), 14. https://doi.org/10.2307/1598947

Dworkin, R. (1977). *Taking rights seriously*. Harvard University Press.

Dworkin, R. (1997). *Law's empire* (10. print). Belknap Press of Harvard Univ. Press.

Dworkin, R. (2011). *Justice for Hedgehogs*. Belknap Press of Harvard University Press.

Ellis, E. (2015). *Kant's Political Theory: Interpretations and Applications*. Penn State University Press.

Engstrom, S. (2009). *The Form of Practical Knowledge: A Study of the Categorical Imperative*. Harvard University Press. https://doi.org/10.4159/9780674053793

Engstrom, S. (2013). Constructivism and practical knowledge. In C. Bagnoli (Ed.), *Constructivism in Ethics* (pp. 133–152). Cambridge University Press. https://doi.org/10.1017/CBO9781139094221.008

Enoch, D. (2006). Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action. *The Philosophical Review*, *115*(2), 169–198. https://doi.org/10.1215/00318108-2005-014

Ereshefsky, M. (2022). Species. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2022). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2022/entries/species/

Etinson, A. (2020). *What's so special about human dignity?* https://doi.org/10.1111/papa.12175

Falk, W. D. (1948). "Ought" and Motivation. *Proceedings of the Aristotelian Society*, *48*, 111–138.

Falk, W. D. (1986). *Ought, reasons, and morality: The collected papers of W. D. Falk* (K. Baier, Ed.). Cornell Univ. Press.

Feinberg, J. (1966). Duties, Rights, and Claims. *American Philosophical Quarterly*, *3*(2), 137–144.

Feinberg, J. (1970). The nature and value of rights. *The Journal of Value Inquiry*, *4*(4), 243–260. https://doi.org/10.1007/BF00137935

Feinberg, J. (1974). The Rights of Animals and Unborn Generations. In W. T. Blackstone (Ed.), *Philosophy & Environmental Crisis* (pp. 43–68). https://philarchive.org/rec/FEITRO-17

Feinberg, J. (1980). *Rights, justice, and the bounds of liberty: Essays in social philosophy*. Princeton University Press.

Feinberg, J. (1992). In Defence of Moral Rights. *Oxford Journal of Legal Studies*, *12*(2), 149–169. https://doi.org/10.1093/ojls/12.2.149

Ferrara, A. (2014). *The democratic horizon: Hyperpluralism and the renewal of political liberalism*. Cambridge University Press.

Ferrero, L. (2009). Constitutivism and the Inescapability of Agency. *Oxford Studies in Metaethics*, *4*, 303–333.

Finlay, S., & Schroeder, M. (2017). Reasons for Action: Internal vs. External. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2017). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2017/entries/reasons-internal-external/

Finlayson, J. G. (2019). *The Habermas-Rawls debate*. Columbia university press.

Finnis, J. (1999). On the Incoherence of Legal Positivism. *Notre Dame Law Review*, *75*, 1597.

Finnis, J. (2011). *Natural Law and Natural Rights* (Second Edition,). Oxford University Press.

Finnis, J. (2020). Natural Law Theories. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2020). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2020/entries/natural-law-theories/

Flikschuh, K. (2008). Reason, Right, and Revolution: Kant and Locke. *Philosophy & Public Affairs*, *36*(4), 375–404.

Flikschuh, K. (2015). Human Rights in Kantian Mode: A Sketch. In R. Cruft, S. M. Liao, & M. Renzo (Eds.), *Philosophical Foundations of Human Rights* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199688623.003.0038

Flikschuh, K. (2017). A Regime of Equal Private Freedom? : Individual Rights and Public Law in Ripstein's Force and Freedom. In S. Kisilevsky & M. J. Stone (Eds.), *Freedom and Force: Essays on Kant's Legal Philosophy* (pp. 55–77). Bloomsbury.

Flynn, J. (2011). Human rights in history and contemporary practice: Source materials for philosophy. In *Philosophical dimensions of human rights: Some contemporary views* (pp. 3–22). Springer.

Foot, P. (1972). Morality as a System of Hypothetical Imperatives. *The Philosophical Review*, *81*(3), 305. https://doi.org/10.2307/2184328

Foot, P. (2003). *Natural goodness* (1. publ. in paperback). Clarendon Press.

Formosa, P. (2008). "All Politics Must Bend Its Knee Before Right": Kant on the Relation of Morals to Politics. *Social Theory and Practice*, *34*(2), 157–181.

Formosa, P. (2013). Is Kant a Moral Constructivist or a Moral Realist? *European Journal of Philosophy*, *21*(2), 170–196. https://doi.org/10.1111/j.1468-0378.2010.00438.x

Fornaroli, G. (2024). Corrective Duties/Corrective Justice. *Philosophy Compass*, *19*(3), e12968. https://doi.org/10.1111/phc3.12968

Forst, R. (1999). The basic right to Justification: Toward a Constructivist Conception of Human Rights. *Constellations*, *6*.

Forst, R. (2001). Towards a Critical Theory of Transnational Justice. *Metaphilosophy*, *32*(1/2), 160–179.

Forst, R. (2002). *Contexts of justice: Political philosophy beyond liberalism and communitarianism*. University of California Press.

Forst, R. (2007). First Things First: Redistribution, Recognition and Justification. *European Journal of Political Theory*, *6*(3), 291–304. https://doi.org/10.1177/1474885107077318

Forst, R. (2010). The Justification of Human Rights and the Basic Right to Justification: A Reflexive Approach. *Ethics*, *120*(4), Article 4. https://doi.org/10.1086/653434

Forst, R. (2011). The Ground of Critique: On the Concept of Human Dignity in Social Orders of Justification. *Philosophy and Social Criticism*, *37*(9), 965–976. https://doi.org/10.1177/0191453711416082

Forst, R. (2012). *The right to justification: Elements of a constructivist theory of justice* (J. Flynn, Trans.; paperback ed). Columbia Univ. Press.

Forst, R. (2013). *Toleration in Conflict: Past and Present* (C. Ciaran, Trans.). Cambridge University Press. https://doi.org/10.1017/CBO9781139051200

Forst, R. (2014). *Justification and critique: Towards a critical theory of politics* (C. Cronin, Trans.; English ed). Polity.

Forst, R. (2016). The Justification of Basic Rights. *Netherlands Journal of Legal Philosophy*, *45*(3), 7–28. https://doi.org/10.5553/NJLP/221307132016045003002

Forst, R. (2017a). *Normativity and power: Analyzing social orders of justification* (C. Ciaran, Trans.). Oxford university press.

Forst, R. (2017b). Political Liberalism: A Kantian View. *Ethics*, *128*(1), 123–144.

Forst, R. (2021). Justification Fundamentalism: A Discourse-Theoretical Interpretation of Scanlon's Contractualism. In M. Stepanians & M. Frauchiger (Eds.), *Reason, Justification, and Contractualism* (pp. 45–58). De Gruyter. https://doi.org/10.1515/9783110733679-006

Forst, R. (2022). The Justification of Trust in Conflict: Conceptual and Normative Groundwork. *ConTrust Working Paper 2,*. https://contrust.uni-frankfurt.de/files/2022/10/ConTrust - WorkingPaper-No2_Forst.pdf. —, "The Meaning(s) of Solidarity."

Forst, R. (2023). The rational critique of social unreason. On critical theory in the Frankfurt tradition. *Constellations*, *30*(4), 395–400. https://doi.org/10.1111/1467-8675.12724

Forst, R. (2024). *The Noumenal republic: Critical constructivism after Kant*. Polity Press.

Francione, G. L. (2006, December 20). The Great Ape Project: Not so Great. *Animal Rights The Abolitionist Approach*. https://www.abolitionistapproach.com/the-great-ape-project-not-so-great/

Frankena, W. K. (1939). The Naturalistic Fallacy. *Mind*, *48*(192), 464–477.

Frankfurt, H. G. (1988). *Importance of What We Care About: Philosophical Essays* (1st ed). Cambridge University Press.

Fraser, N. (1990). Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy. *Social Text*, *25/26*, 56–80. https://doi.org/10.2307/466240

Freedman, R. (2014). UN Immunity or Impunity? A Human Rights Based Challenge. *European Journal of International Law*, *25*(1), 239–254. https://doi.org/10.1093/ejil/cht082

Freeman, S. (2000). Deliberative Democracy: A Sympathetic Comment. *Philosophy & Public Affairs*, *29*(4), 371–418. https://doi.org/10.1111/j.1088-4963.2000.00371.x

Freeman, S. (2006). Frontiers of justice: The capabilities approach vs. contractarianism. *Texas Law Review*, *85*(2), 385.

Freeman, S. (2011). Problems with some consequentialist arguments for basic rights. In *Problems with some consequentialist arguments for basic rights* (pp. 107–128). De Gruyter. https://doi.org/10.1515/9783110263886.107

Frege, F. L. G. (1903). *Grundgesetze der Arithmetik Vol. (Band 2)*. Jena: Verlag Hermann Pohle.

Freyenhagen, F. (2011). Taking reasonable pluralism seriously: An internal critique of political liberalism. *Politics, Philosophy & Economics*, *10*(3), 323–342. https://doi.org/10.1177/1470594X10388381

Frydrych, D. (2019). What Is the Will Theory of Rights? *Ratio Juris*, *32*(4), 455–472. https://doi.org/10.1111/raju.12259

Fulda, H. F. (1997). Kants Postulat des Öffentlichen Rechts (RL § 42). In B. S. Byrd (Ed.), *Themenschwerpunkt: 200 Jahre Kants Metaphysik der Sitten* (pp. 267–290). Duncker & Humblot. https://doi.org/10.11588/HEIDOK.00017439

Fuller, L. L. (1978). *The morality of law* (Rev. ed., 15. print). Yale Univ. Press.

Garner, R. T. (1990). On the genuine queerness of moral properties and facts. *Australasian Journal of Philosophy*, *68*(2), 137–146. https://doi.org/10.1080/00048409012344161

Gaus, G. (2010). *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge University Press. https://doi.org/10.1017/CBO9780511780844

George, R. P. (2001). Reason, Freedom, and the Rule of Law: Their Significance in the Natural Law Tradition. *American Journal of Jurisprudence*, *46*(1), 249.

George, R. P. (2004). *In defense of natural law* (Reprinted). Oxford Univ. Press.

Gert, B., & Gert, J. (2020). The Definition of Morality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2020). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2020/entries/morality-definition/>.

Gert, J. (2002). Korsgaard's Private-Reasons Argument. *Philosophy and Phenomenological Research*, *64*(2), 303–324. https://doi.org/10.1111/j.1933-1592.2002.tb00003.x

Gewirth, A. (1980). *Reason and Morality*. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/R/bo25842059.html

Gewirth, A. (1982). *Human rights: Essays on justification and applications*. University of Chicago Press.

Gewirth, A. (1984). The Epistemology of Human Rights. *Social Philosophy and Policy*, *1*(2), 1–24. https://doi.org/10.1017/S0265052500003836

Gibbard, A. (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press.

Gilabert, P. (2013). The Capability Approach and the Debate Between Humanist and Political Perspectives on Human Rights. A Critical Survey. *Human Rights Review*, *14*(4), 299–325. https://doi.org/10.1007/s12142-013-0269-z

Gilbert, M. (2018). *Rights and demands: A foundational inquiry*. Oxford university press.

Glüer, K., Wikforss, Å., & Ganapini, M. (2023). The Normativity of Meaning and Content. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2023). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2023/entries/meaning-normativity/

Goldin, C. D. (1990). *Understanding the gender gap: An economic history of American women*. Oxford university press.

Gosepath, S. (2021). Equality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2021/entries/equality/

Gray, J. (2000). *Two Faces of Liberalism*. The New Press.

Green, L., & Adams, T. (2019). Legal Positivism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2019). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2019/entries/legal-positivism/

Griffin, J. (2008). *On human rights*. Oxford University Press.

Griffin, J. (2010). Human Rights and the Autonomy of International Law. In S. Besson & J. Tasioulas (Eds.), *The philosophy of international law* (pp. 339–355). Oxford University Press.

Grube, D.-M. (1998). *Realism, Foundationalism, and Constructivism: A Philosopher's Bermuda-Triangle? 40*(1), 107–126. https://doi.org/10.1515/nzst.1998.40.1.107

Guillaume, X., & Forst, R. (2012). The Right to Justification: Towards a Critical Theory of Justice and Democracy. An Interview with Rainer Forst. In G. Browning, R. Prokhovnik, & M. Dimova-Cookson (Eds.), *Dialogues with Contemporary Political Theorists* (pp. 105–120). Palgrave Macmillan UK. https://doi.org/10.1057/9781137271297_7

Guyer, P. (2005). *Kant's System of Nature and Freedom: Selected Essays*. Oxford University Press.

Haakonssen, K. (1996). *Natural Law and Moral Philosophy: From Grotius to the Scottish Enlightenment*. New York, NY, USA: Cambridge University Press.

Haase, M. (2012). Three Forms of the First Person Plural. In G. Abel & J. Conant (Eds.), *Rethinking Epistemology: Volume 2* (pp. 229–256). De Gruyter. https://doi.org/10.1515/9783110277944.229

Haase, M., & Mayr, E. (2019). Varieties of constitutivism. *Philosophical Explorations*, *22*(2), 95–97. https://doi.org/10.1080/13869795.2019.1601754

Habermas, J. (1979). *Communication and the evolution of society*. Beacon Press.

Habermas, J. (1985). *The Theory of Communicative Action* (Vol. 1). Beacon Press.

Habermas, J. (1989). Justice and solidarity: On the discussion concerning Stage 6. In T. E. Wren (Ed.), *The moral domain: Essays in the ongoing discussion between philosophy and the social sciences* (p. 251). The MIT Press.

Habermas, J. (1990). *Moral consciousness and communicative action*. Polity.

Habermas, J. (1993). *Justification and application: Remarks on discourse ethics*. MIT Press.

Habermas, J. (1995). Reconciliation Through the Public use of Reason: Remarks on John Rawls's Political Liberalism. *The Journal of Philosophy*, *92*(3), 109–131. https://doi.org/10.2307/2940842

Habermas, J. (1996). *Between facts and norms: Contributions to a discourse theory of law and democracy* (W. Rehg, Trans.). MIT Press.

Habermas, J. (1998). *The inclusion of the other: Studies in political theory* (C. Cronin & P. De Greiff, Eds.). MIT PRess.

Habermas, J. (1999). Between Facts and Norms: An Author's Reflections. *Denver Law Review*, *76*(4), 937.

Habermas, J. (2008). *Between naturalism and religion: Philosophical essays*. Polity Press.

Habermas, J. (2011). "Reasonable" versus "True," or the Morality of Worldviews. In J. G. Finlayson (Ed.), *Habermas and Rawls*. Routledge.

Hammerton, M. (2019). Distinguishing Agent-Relativity from Agent-Neutrality. *Australasian Journal of Philosophy*, *97*(2), 239–250. https://doi.org/10.1080/00048402.2018.1477166

Hanisch, C. (2018). KANT'S JURISPRUDENCE: Beyond natural law and legal positivism. *Pólemos. Materiali di filosofia e critica sociale*, *1*, 95–119.

Hannum, H. (2016). *Reinvigorating Human Rights for the Twenty-First Century* (SSRN Scholarly Paper No. 2832205). Social Science Research Network. https://papers.ssrn.com/abstract=2832205

Hare, R. M. (1963a). *Freedom and reason*. Clarendon Press.

Hare, R. M. (1963b). *The Language of Morals*. Oxford University PressOxford. https://doi.org/10.1093/0198810776.001.0001

Hare, R. M. (1992). *Moral thinking: Its levels, method, and point* (7th impr). Clarendon Pr.

Harman, G. (1986). *Change in view: Principles of reasoning*. MIT Press.

Harman, G. (2004). Practical Aspects of Theoretical reasoning. In A. R. Mele & P. Rawling (Eds.), *The Oxford Handbook of Rationality* (pp. 45–55). Oxford University Press. https://doi.org/10.1093/0195145399.003.0003

Harris, E. by J. (Ed.). (2001). *Bioethics*. Oxford University Press.

Harsanyi, J. C. (1977). *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511571756

Hart, H. L. A. (1955). Are There Any Natural Rights? *The Philosophical Review*, *64*(2), 175–191. https://doi.org/10.2307/2182586

Hart, H. L. A. (1958). Positivism and the Separation of Law and Morals. *Harvard Law Review*, *71*(4), 593. https://doi.org/10.2307/1338225

Hart, H. L. A. (1961). *The concept of law*. Clarendon Press.

Hart, H. L. A. (1983). *Essays in jurisprudence and philosophy*. Clarendon Press.

Hasan, A., & Fumerton, R. (2022). Foundationalist Theories of Epistemic Justification. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2022).

Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2022/entries/justep-foundational/

Hayward, T. (2013). On Prepositional Duties. *Ethics*, *123*(2), 264–291. https://doi.org/10.1086/668706

Heath, J. (2014). Rebooting discourse ethics. *Philosophy & Social Criticism*, *40*(9), 829–866. https://doi.org/10.1177/0191453714545340

Herman, B. (1993). *The practice of moral judgment*. Harvard Univ. Press.

Herstein, O. (2023). Legal Rights. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2023). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2023/entries/legal-rights/

Hessler, K. (2017). Theory, politics, and practice: Methodological pluralism in the philosophy of human rights. In *Moral and Political Conceptions of Human Rights: Implications for Theory and Practice* (pp. 15–32). Cambridge University Press.

Heyd, D. (2024). Supererogation. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Spring 2024). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2024/entries/supererogation/

Hobbes, T. (1651). *Leviathan* (R. Tuck, Ed.; [1996]). Cambridge University Press.

Hohfeld, W. N. (with Robarts - University of Toronto). (1920). *Fundamental legal conceptions as applied in judicial reasoning: And other legal essays* (walter W. Cook, Ed.). New Haven : Yale University Press. http://archive.org/details/fundamentallegal00hohfuoft

Honneth, A. (1995). *The struggle for recognition: The moral grammar of social conflicts* (J. Anderson, Trans.). Polity Press.

Horkheimer, M. (2013). *Eclipse of reason*. Bloomsbury.

Horkheimer, M., & Adorno, T. W. (2002). *Dialectic of enlightenment: Philosophical fragments* (G. Schmid Noerr, Ed.; Nachdr.). Stanford Univ. Press.

Howse, R., & Teitel, R. (2010). Global Justice, Poverty and the International Economic Order. In S. Besson & J. Tasioulas (Eds.), *The Philosophy of International Law*. Oxford University Press.

Iglesias, T. (2001). Bedrock Truths and the Dignity of the Individual. *Logos: A Journal of Catholic Thought and Culture*, *4*(1), 114–134.

Ignatieff, M. (2003). *Human rights as politics and idolatry* (A. Gutmann, Ed.; 3. print., and 1. paperback print). Princeton University Press.

Johnson, R., & Cureton, A. (2024). Kant's Moral Philosophy. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2024). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2024/entries/kant-moral/

Jones, P. (1994). *Rights* (1. publ). Macmillan.

Jonker, J. (2020). Directed Duties and Moral Repair. *Philosopher's Imprint*, *20*(23). http://hdl.handle.net/2027/spo.3521354.0020.023

Joseph, H. W. B. (1933). *Some Problems in Ethics*. Oxford University Press.

Joyce, R. (2001). *The Myth of Morality* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511487101

Kagan, S. (2016). What's Wrong with Speciesism? (Society for Applied Philosophy Annual Lecture 2015). *Journal of Applied Philosophy*, *33*(1), 1–21. https://doi.org/10.1111/japp.12164

Kamenka, E., & Tay, A. E.-S. (1984). Philosophy and Human Rights: A Survey and Select Annotated Bibliography of Recent English-Language Literature. *ARSP: Archiv Für*

Rechts- Und Sozialphilosophie / Archives for Philosophy of Law and Social Philosophy, *70*(3), 291–341.

Kamm, F. M. (2006). *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. Oxford University Press.

Kelsen, H. (1967). *Pure Theory of Law*. University of California Press. https://doi.org/10.1525/9780520312296

Kennedy, D. (2004). *The Dark Sides of Virtue: Reassessing International Humanitarianism*. Princeton University Press. https://www.jstor.org/stable/j.ctt7rsjc

Kinghorn, K. (2021). *The Nature of Desert Claims: Rethinking What it Means to Get One's Due*. Cambridge University Press. https://doi.org/10.1017/9781108955546

Kisilevsky, S., & Stone, M. J. (Eds.). (2019). *Freedom and force: Essays on Kant's legal philosophy* (Paperback edition). Hart.

Klein, P. D. (1999). Human Knowledge and the Infinite Regress of Reasons. *Philosophical Perspectives*, *13*(Epistemology), 297–325.

Kolnai, A. (1961). Deliberation Is of Ends. *Proceedings of the Aristotelian Society*, *62*, 195–218.

Koons, R. (2022). Defeasible Reasoning. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2022). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2022/entries/reasoning-defeasible/

Korsgaard, C. M. (1983). Two Distinctions in Goodness. *The Philosophical Review*, *92*(2), 169. https://doi.org/10.2307/2184924

Korsgaard, C. M. (1986). Skepticism About Practical Reason. *Journal of Philosophy*, *83*(1), 5–25.

Korsgaard, C. M. (1993). The Reasons we can Share: An Attack on the Distinction between Agent-Relative and Agent-Neutral Values. *Social Philosophy and Policy*, *10*(1), 24–51. https://doi.org/10.1017/S0265052500004003

Korsgaard, C. M. (1996a). *Creating the kingdom of ends*. Cambridge University Press.

Korsgaard, C. M. (1996b). *The Sources of Normativity*. Cambridge University Press. https://doi.org/10.1017/CBO9780511554476

Korsgaard, C. M. (1999). Self-Constitution in the Ethics of Plato and Kant. *The Journal of Ethics*, *3*(1), 1–29.

Korsgaard, C. M. (2004). Fellow creatures: Kantian ethics and our duties to animals. *The Tanner Lectures on Human Values*.

Korsgaard, C. M. (2008). *The Constitution of Agency: Essays on Practical Reason and Moral Psychology* (1st ed.). Oxford University PressOxford. https://doi.org/10.1093/acprof:oso/9780199552733.001.0001

Korsgaard, C. M. (2009). *Self-Constitution: Agency, Identity, and Integrity*. Oxford University Press.

Korsgaard, C. M. (2018). *Fellow creatures: Our obligations to the other animals*. Oxford university press.

Korsgaard, C. M. (2019). Constitutivism and the virtues. *Philosophical Explorations*, *22*(2), 98–116. https://doi.org/10.1080/13869795.2019.1599050

Korsgaard, C. M. (2021). Valuing Our Humanity. In R. Dean & O. Sensen (Eds.), *Respect* (pp. 171–191). Oxford University Press. https://doi.org/10.1093/oso/9780198824930.003.0009

Koskenniemi, M. (2001). *The Gentle Civilizer of Nations: The Rise and Fall of International Law 1870–1960*. Cambridge University Press. https://doi.org/10.1017/CBO9780511494222

Kramer, M. H. (1999). *In defense of legal positivism: Law without trimmings*. Oxford University Press.

Kramer, M. H. (2000). Rights Without Trimmings. In M. H. Kramer, N. E. Simmonds, & H. Steiner (Eds.), *A Debate Over Rights*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198298991.001.0001

Kramer, M. H. (2013). Some Doubts about Alternatives to the Interest Theory of Rights. *Ethics*, *123*(2), 245–263. https://doi.org/10.1086/668705

Kramer, M. H., & Steiner, H. (2007). Theories of Rights: Is There a Third Way? *Oxford Journal of Legal Studies*, *27*(2), 281–310. https://doi.org/10.1093/ojls/gqi039

Kramm, M. (2020). When a River Becomes a Person. *Journal of Human Development and Capabilities*, *21*(4), 307–319. https://doi.org/10.1080/19452829.2020.1801610

Kurki, V. A. J. (2019). Rights and Demands. *Analysis*, *79*(3), 530–538. https://doi.org/10.1093/analys/anz036

Laitinen, A. (2008). *Strong Evaluation without Moral Sources: On Charles Taylor's Philosophical Anthropology and Ethics*. Walter de Gruyter. https://doi.org/10.1515/9783110211900

Larmore, C. E. (Ed.). (1996). *The Morals of Modernity*. Cambridge University Press. https://www.cambridge.org/core/books/morals-of-modernity/index/C1A67EB378BCC30805EE78C8EE552DB4

Larmore, C. E. (1999). The Moral Basis of Political Liberalism. *The Journal of Philosophy*, *96*(12), 599–625. https://doi.org/10.2307/2564695

Larmore, C. E. (2020). *What is political philosophy?* Princeton University press.

LeBar, M. (2001). Korsgaard, Wittgenstein, and the Mafioso. *The Southern Journal of Philosophy*, *39*(2), 261–271. https://doi.org/10.1111/j.2041-6962.2001.tb01817.x

LeBar, M. (2008). Aristotelian Constructivism. *Social Philosophy and Policy*, *25*(1), 182–213. https://doi.org/10.1017/S0265052508080072

LeBar, M. (2013). Constructivism and particularism. In C. Bagnoli (Ed.), *Constructivism in Ethics* (1st ed., pp. 183–200). Cambridge University Press. https://doi.org/10.1017/CBO9781139094221.010

Lehrer, K. (1999). Self-Trust: A Study of Reason, Knowledge and Autonomy. *Philosophy and Phenomenological Research*, *59*(4), 1039. https://doi.org/10.2307/2653569

Leiter, B. (2003). Beyond the Hart/Dworkin Debate: The Methodology Problem in Jurisprudence. *The American Journal of Jurisprudence*, *48*(1), 17–51. https://doi.org/10.1093/ajj/48.1.17

Leiter, B. (2007). *Naturalizing jurisprudence: Essays on american legal realism and naturalism in legal philosophy*. Oxford University press.

Lenman, J. (2010). Humean Constructivism in Moral Theory. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics, Volume 5*. Oxford University Press.

Li, Z. (2016). The public conception of morality in John Rawls' political liberalism. *Ethics & Global Politics*, *9*(1), 28679. https://doi.org/10.3402/egp.v9.28679

Liao, S. M., & Etinson, A. (2012). Political and naturalistic conceptions of human rights: A false polemic? *Journal of Moral Philosophy*, *9*(3), 327–352. https://doi.org/10.1163/17455243-00903008

Liberto, H. (2014). The Moral Specification of Rights: A Restricted Account. *Law and Philosophy*, *33*(2), 175–206.

Locke, J. (1689). A Letter Concerning Toleration. In S. Cahn (Ed.), *Classics of Modern Political Theory [1993]*. Oxford University Press.

Locke, J. (1690). *An Essay Concerning Human Understanding* (P. H. Nidditch, Ed.; Clarendon 1st edition [1975]). Clarendon Press.

MacCormick, J. P. (2018). *Reading Machiavelli: Scandalous Books, Suspect Engagements, and the Virtue of Populist Politics*. Princeton University Press.

MacCormick, N. (1977). Rights in Legislation. In *Law, Morality and Society: Essays in Honour of H.L.A Hart. Hacker, and Joseph Raz (eds)* (pp. 189–209). Clarendon Press.

MacCormick, N. (1982). *Children's Rights: A Test Case for Theories of Right*.

MacCormick, N. (2003). *Legal reasoning and legal theory* (Reprinted). Clarendon Press.

MacIntyre, A. C. (1981). *After virtue: A study in moral theory*. University of Notre Dame Press.

Mackie, J. L. (1977). *Ethics: Inventing right and wrong*. Penguin.

Macklin, R. (2003). Dignity is a useless concept. *BMJ : British Medical Journal*, *327*(7429), 1419. https://doi.org/10.1136/bmj.327.7429.1419

Maliks, R., & Schaffer, J. K. (2017a). *Moral and political conceptions of human rights: Implications for theory and practice*. Cambridge University Press.

Maliks, R., & Schaffer, J. K. (2017b). *Moral and political conceptions of human rights: Implications for theory and practice*. Cambridge University Press.

Mandle, J., & Reidy, D. A. (Eds.). (2014). *The Cambridge Rawls Lexicon*. Cambridge University Press. https://doi.org/10.1017/CBO9781139026741

Mayer, S. (2017). Human Rights Solidarity Moral or Political? *Moral and Political Conceptions of Human Rights: Implications for Theory and Practice*, 157.

McCarthy, T. (1994). Kantian Constructivism and Reconstructivism: Rawls and Habermas in Dialogue. *Ethics*, *105*(1), 44–63.

McDowell, J. H. (1996). *Mind and world: With a new introduction* (1st Harvard University Press paperback ed). Harvard University Press.

McDowell, J. H. (2002). *Mind, value, and reality* (2. print). Harvard Univ. Press.

McMahon, C. (2002). Why There is No Issue Between Habermas and Rawls. *Journal of Philosophy*, *99*(3), 111. https://doi.org/10.2307/3655639

McNaughton, D., & Rawling, P. (1991). Agent-Relativity and the Doing-Happening Distinction. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *63*(2), 167–185.

Mill, J. S. (1861). *Utilitarianism* (2014 ed). Cambridge University Press. https://doi.org/10.1017/CBO9781139923927

Millgram, E. (2005). *Ethics Done Right: Practical Reasoning as a Foundation for Moral Theory*. Cambridge University Press. https://doi.org/10.1017/CBO9780511610615

Millgram, E. (2008). Specificationism. In J. E. Adler & L. J. Rips (Eds.), *Reasoning: Studies of Human Inference and its Foundations* (pp. 731–747).

Millgram, E. (2022). Bounded agency. In L. Ferrero (Ed.), *The Routledge Handbook of Philosophy of Agency*. Routledge.

Moore, G. E. (1903). *Principia Ethica* (T. Baldwin, Ed.). Dover Publications.

Moore, G. E. (2004). *Principia ethica*. Dover Publications.

Morelli, J. (2023). Kantian Constructivism and Practical Reasoning in Clinical Bioethics. *Qeios*. https://doi.org/10.32388/3HFS7Y.2

Moreso, J. J. (2012). Ways of Solving Conflicts of Constitutional Rights: Proportionalism and Specificationism. *Ratio Juris*, *25*(1), 31–46. https://doi.org/10.1111/j.1467-9337.2011.00501.x

Moyn, S. (2012). *The last utopia: Human rights in history*. Harvard University Press.

Mulholland, L. A. (1990). *Kant's System of Rights* (p. 434 Pages). Columbia University Press.

Müller, A. (2020). *Constructing practical reasons* (First edition). Oxford University Press.

Müller, L. K. (2017). Rawls's Relational Conception of Human Rights. In J. K. Schaffer & R. Maliks (Eds.), *Moral and Political Conceptions of Human Rights: Implications for Theory and Practice* (pp. 58–76). Cambridge University Press. https://doi.org/10.1017/9781316650134.005

Murdoch, I. (1970). *The Sovereignty of Good*. Routledge.

Murphy, M. (2019). The Natural Law Tradition in Ethics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2019). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2019/entries/natural-law-ethics/

Mutua, M. W. (2000). Politics and Human Rights: An Essential Symbiosis. In M. Byers (Ed.), *The Role of Law in International Politics* (p. 149). Oxford University Press. https://digitalcommons.law.buffalo.edu/book_sections/206

Nagel, T. (1970). *The possibility of altruism*. Princeton Univ. Press.

Nagel, T. (1979). *Mortal questions*. Cambridge University Press.

Nagel, T. (1986). *The view from nowhere*. Oxford University Press.

Nagel, T. (1995). Personal Rights and Public Space. *Philosophy & Public Affairs*, *24*(2), 83–107. https://doi.org/10.1111/j.1088-4963.1995.tb00024.x

National Academy of Sciences US. (2010). *In the Light of Evolution: Volume IV: The Human Condition* (J. C. Avise & F. J. Ayala, Eds.). National Academies Press (US). http://www.ncbi.nlm.nih.gov/books/NBK210014/

Nickel, J. W. (1979). Review of Human Rights by Kamenka and Tay. *California Law Review*, *67*(5), 1232. https://doi.org/10.2307/3480014

Nickel, J. W. (2006). Are Human Rights Mainly Implemented by Intervention? In *Rawls's Law of Peoples* (pp. 263–277). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470776612.ch15

Nickel, J. W. (2007). *Making sense of human rights* (2nd ed). Blackwell Pub.

Nickel, J. W., & Etinson, A. (2024). Human Rights. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2024). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2024/entries/rights-human/

Nino, C. S. (1980). Dworkin and Legal Positivism. *Mind*, *89*(356), 519–543.

Nino, C. S. (1991). *The Ethics of Human Rights*. Oxford University Press.

Nozick, R. (1974). *Anarchy, state, and utopia*. Basic Books.

Nozick, R. (1994). *The Nature of Rationality*. Princeton University Press.

Nussbaum, M. C. (1986). *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge University Press. https://doi.org/10.1017/CBO9780511817915

Nussbaum, M. C. (1997). *Capabilities and Human Rights*. *66 Fordham L. Rev. 273*. https://doi.org/Available%2520at:%2520https://ir.lawnet.fordham.edu/flr/vol66/iss2/2

Nussbaum, M. C. (2000). *Women and Human Development: The Capabilities Approach*. Cambridge University Press. https://doi.org/10.1017/CBO9780511841286

Nussbaum, M. C. (2007). *Frontiers of Justice: Disability, Nationality, Species Membership*. Harvard University Press. https://doi.org/10.2307/j.ctv1c7zftw

Nussbaum, M. C. (2013). *Creating Capabilities: The Human Development Approach*. Belknap Press.

Nwaneri, N. V. (2017). *Human Rights and Global Justice: A Normative Critique of Some Rawlsian Approaches.* Loyola eCommons.

O'Neill, O. (1986). *Faces of Hunger: An Essay on Poverty, Justice, and Development*. G. Allen & Unwin.

O'Neill, O. (1987). [Review of *Impartial Reason*, by S. Darwall]. *The Journal of Philosophy*, *83*(1), 60–64. https://doi.org/10.2307/2026467

O'Neill, O. (1988). Constructivisms in Ethics. *Proceedings of the Aristotelian Society*, *89*, 1–17.

O'Neill, O. (1989). *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge University Press. https://doi.org/10.1017/CBO9781139173773

O'Neill, O. (1992). Children's Rights and Children's Lives. *International Journal of Law, Policy and the Family*, *6*(1), 24–42. https://doi.org/10.1093/lawfam/6.1.24

O'Neill, O. (1996). *Towards justice and virtue: A constructive account of practical reasoning* (Transferred to digital print). Cambridge University Press.

O'Neill, O. (1997). Environmental Values, Anthropocentrism and Speciesism. *Environmental Values*, *6*(2), 127–142. https://doi.org/10.3197/096327197776679121

O'Neill, O. (2000). *Bounds of Justice*. Cambridge University Press. https://doi.org/10.1017/CBO9780511605734

O'Neill, O. (2001). Agents of Justice. *Metaphilosophy*, *32*(1/2), 180–195.

O'Neill, O. (2003). Constructivism VS. Contractualism. *Ratio*, *16*(4), 319.

O'Neill, O. (2005). The Dark Side of Human Rights. *International Affairs (Royal Institute of International Affairs 1944-)*, *81*(2), 427–439.

O'Neill, O. (2015a). *Constructing Authorities: Reason, Politics and Interpretation in Kant's Philosophy* (1st ed.). Cambridge University Press. https://doi.org/10.1017/cbo9781316337141

O'Neill, O. (Ed.). (2015b). Kant and the Social Contract tradition. In *Constructing Authorities: Reason, Politics and Interpretation in Kant's Philosophy* (pp. 170–185). Cambridge University Press. https://doi.org/10.1017/CBO9781316337141.012

O'Neill, O. (2016). *Justice across Boundaries: Whose Obligations?* Cambridge University Press.

Onuf, N. G. (2013). *World of our making: Rules and rule in social theory and international relations* (reissued). Routledge.

Palmquist, S. (2022). Methods of Applying Kant's Ethics—A Reply to Jacopo Morelli's Reflections on Clinical Bioethics. *Qeios*. https://doi.org/10.32388/66YWKP

Parfit, D. (1984). *Reasons and persons* (Repr. with corr). Clarendon Pr.

Parfit, D. (2011). *On What Matters: Volume One*. Oxford University Press.

Parfit, D. (2017). *On What Matters: Volume Three*. Oxford University Press.

Parkin, J. (2007). *Taming the Leviathan: The Reception of the Political and Religious Ideas of Thomas Hobbes in England 1640–1700* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511720499

Passerin d'Entrèves, A. (1970). *Natural Law: An Introduction to Legal Philosophy*. Hutchinson University Library.

Pavlakos, G. (2017). *Why is willing irrelevant to the grounding on any obligation? Remarks on A. Ripstein's conception of omni-lateral willing* (S. Kisilevsky & M. J. Stone, Eds.; pp. 113–127). Hart Publishing. http://www.hartpub.co.uk/BookDetails.aspx?ISBN=9781849463164

Pelenc, J., Lompo, M. K., Ballet, J., & Dubois, J.-L. (2016). Sustainable human development and the capability approach: Integrating environment, responsibility and collective agency. In *The Capability Approach and Sustainability* (pp. 76–93). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315540832-6/sustainable-human-development-capability-approach-integrating-environment-responsibility-collective-agency-j%C3%A9r%C3%B4me-pelenc-minkieba-kevin-lompo-j%C3%A9r%C3%B4me-ballet-jean-luc-dubois

Perry, M. J. (2000). *The Idea of Human Rights: Four Inquiries*. Oxford University Press.

Phillips, A. (2021). *Unconditional equals*. Princeton University Press.

Pinker, S. (2012). *The Better Angels of Our Nature: Why Violence Has Declined* (Reprint edizione). Penguin Books.

Piper, A. M. S. (2012). Kant's Self-Legislation Procedure Reconsidered. *Kant Studies Online*, *2012*(1), 203–277.

Plantinga, A. (2000). *Warranted Christian belief*. Oxford Univ. Press.

Pogge, T. W. (1989). *Realizing Rawls*. Cornell University Press.

Pogge, T. W. (1992). O'neill on Rights and Duties. *Grazer Philosophische Studien*, *43*(1), Article 1. https://doi.org/10.1163/18756735-04301014

Pogge, T. W. (1995). How Should Human Rights be Conceived? *Jahrbuch Für Recht Und Ethik / Annual Review of Law and Ethics*, *3*, 103–120.

Pogge, T. W. (2008). *World poverty and human rights: Cosmopolitan responsibilities and reforms* (2nd ed). Polity.

Pogge, T. W. (2014). Are We Violating the Human Rights of the World's Poor? *Yale Human Rights and Development Law Journal*. https://openyls.law.yale.edu/handle/20.500.13051/5760

Pöschl, V. (1969). *Der Begriff der Würde im antiken Rom und später*. Winter.

Pufendorf, S. (1717). *On the duty of man and citizen according to natural law* (J. Tully & M. Silverthorne, Eds.; Cambridge Texts in the History of Political Thought [1991]). Cambridge University Press.

Putnam, H. (1982). Beyond the Fact-Value Dichotomy. *Crítica: Revista Hispanoamericana de Filosofía*, *14*(41), 3–12.

Quinn, G., & Arstein-Kerslake, A. (2012). Restoring the 'human' in 'human rights': Personhood and doctrinal innovation in the UN disability convention. In C. Gearty & C. Douzinas (Eds.), *The Cambridge Companion to Human Rights Law* (pp. 36–55). Cambridge University Press. https://doi.org/10.1017/CCO9781139060875.005

Quinn, W. (1993). *Morality and Action*. Cambridge University Press. https://doi.org/10.1017/CBO9781139172677

Quong, J. (2011). *Liberalism without perfection*. Oxford Univ. Press. https://doi.org/10.1093/acprof:oso/9780199594870.001.0001

Rabinowicz, W., & Rønnow-Rasmussen, T. (2000). A Distinction in Value: Intrinsic and for its Own Sake. *Proceedings of the Aristotelian Society*, *100*(1), 33–51. https://doi.org/10.1111/j.0066-7372.2003.00002.x

Railton, P. (1984). Alienation, Consequentialism, and the Demands of Morality. *Philosophy and Public Affairs*, *13*(2), 134–171.

Railton, P. (1986). Moral Realism. *Philosophical Review*, *95*(2), 163–207. https://doi.org/10.2307/2185589

Rauscher, F. (2015). *Naturalism and Realism in Kant's Ethics* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781316105252

Rawls, J. (1955). Two Concepts of Rules. *The Philosophical Review*, *64*(1), 3–32. https://doi.org/10.2307/2182230

Rawls, J. (1971). *A Theory of Justice*. Harvard University Press. https://doi.org/10.2307/j.ctvjf9z6v

Rawls, J. (1980). Rational and Full Autonomy: Kantian Constructivism in Moral Theory. *The Journal of Philosophy*, *77*(9), 515–572. https://doi.org/10.5840/jphil198077936

Rawls, J. (1985). Justice as Fairness: Political Not Metaphysical. *Philosophy and Public Affair*, *14*(3), 223–251. https://doi.org/10.1007/978-1-349-21763-2_10

Rawls, J. (1993). *Political Liberalism*. Columbia University Press.

Rawls, J. (1995). Political Liberalism: Reply to Habermas. *The Journal of Philosophy*, *92*(3), 132–180. https://doi.org/10.2307/2940843

Rawls, J. (1997). The Idea of Public Reason Revisited. *University of Chicago Law Review*, *64*(3), 765–807.

Rawls, J. (1999). *The Law of Peoples*. Harvard University Press.

Rawls, J. (2000). *Lectures on the History of Moral Philosophy* (B. Herman, Ed.; Illustrated Edition). Harvard University Press.

Rawls, J. (2001). *Justice as Fairness: A Restatement* (E. Kelly, Ed.). Harvard University Press. https://doi.org/10.2307/j.ctv31xf5v0

Raz, J. (1970). *The concept of a legal system: An introduction to the theory of legal system*. Clarendon Press.

Raz, J. (1979). *The authority of law: Essays on law and morality*. Clarendon Press ; Oxford University Press.

Raz, J. (1984). On the Nature of Rights. *Mind*, *XCIII*(370), 194–214. https://doi.org/10.1093/mind/XCIII.370.194

Raz, J. (1985). Authority, Law and Morality. *The Monist*, *68*(3), 295–324. https://doi.org/10.5840/monist198568335

Raz, J. (1986). *The Morality of Freedom*. Oxford University Press.

Raz, J. (1999). Explaining Normativity: On Rationality and the Justification of Reason. *Ratio*, *12*(4), 354–379. https://doi.org/10.1111/1467-9329.00099

Raz, J. (2003a). About Morality and the Nature of Law. *American Journal of Jurisprudence, Vol. 48, p. 1, 2003*. https://scholarship.law.columbia.edu/faculty_scholarship/1497

Raz, J. (2003b). Numbers, With and Without Contractualism. *Ratio*, *16*(4), 346–367. https://doi.org/10.1046/j.1467-9329.2003.00228.x

Raz, J. (2009). *Between authority and interpretation: On the theory of law and practical reason*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199562688.001.0001

Raz, J. (2010a). Human Rights Without Foundations. In S. Besson & J. Tasioulas (Eds.), *The Philosophy of International Law*. Oxford University Press.

Raz, J. (2010b). Reason, Reasons and Normativity. *Oxford Studies in Metaethics, Vol. 5, Russ Shafer-Landau (Ed.), 2010*. https://scholarship.law.columbia.edu/faculty_scholarship/1496

Raz, J. (2011). *From normativity to responsibility*. Oxford University Press.

Raz, J. (2022). *The roots of normativity* (U. Heuer, Ed.). Oxford university press.

Reath, A. (2013). Formal approaches to Kant's formula of humanity. In M. Timmons & S. Baiasu (Eds.), *Kant on Practical Justification: Interpretive Essays* (pp. 201–228). Oxford University Press. https://books.google.com/books?hl=en&lr=&id=imBpAgAAQBAJ&oi=fnd&pg=PA201&dq=Reath,+Andrews.+2013.+Formal+approaches+to+Kant%E2%80%99s+Formula+of+Humanity.+In+Kant+on+practical+justification,+ed.+Mark+Timmons+and+Sorin+Baiasu,+201%E2%80%93228.+Oxford:+Oxford+University+Press.&ots=gz7p6lwFfZ&sig=JRZZaM9LpuoCH7AxFR8AWr-nToI

Reath, A., & Timmermann, J. (Eds.). (2010). *Kant's "Critique of Practical Reason": A Critical Guide*. Cambridge University Press. https://doi.org/10.1017/CBO9780511770869

Rettig, C. (2023). Rights and Practical Reasoning: A Practical View on the Specificationism Vs Generalism Debate. *Journal of Value Inquiry*, *1*(1), 1–15. https://doi.org/10.1007/s10790-023-09946-5

Richardson, H. S. (1990). Specifying Norms as a Way to Resolve Concrete Ethical Problems. *Philosophy and Public Affairs*, *19*(4), 279–310.

Richardson, H. S. (1994). *Practical Reasoning about Final Ends*. Cambridge University Press. https://doi.org/10.1017/CBO9781139174275

Richardson, H. S. (2013). Revising moral norms: Pragmatism and the problem of perspicuous description. In C. Bagnoli (Ed.), *Constructivism in Ethics* (pp. 221–242). Cambridge University Press. https://doi.org/10.1017/CBO9781139094221.012

Richardson, H. S. (2018). *Articulating the Moral Community: Toward a Constructive Ethical Pragmatism*. Oxford University Press.

Ridge, M. (2023). Reasons for Action: Agent-Neutral vs. Agent-Relative. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Spring 2023). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2023/entries/reasons-agent/

Ripstein, A. (2009). *Force and Freedom: Kant's Legal and Political Philosophy*. Harvard University Press.

Ripstein, A. (2015). Kant and the Circumstances of Justice. In *Kant's Political Theory* (pp. 42–73). Penn State University Press. https://www.degruyterbrill.com/document/doi/10.1515/9780271059860-004/html

Robson, G. (2021). Magistrates, Mobs, and Moral Disagreement: Countering the Actual Disagreement Challenge to Moral Realism. *Canadian Journal of Philosophy*, *51*(6), 416–435. https://doi.org/10.1017/can.2021.40

Rolston III, H. (1994). *Conserving Natural Value* (p. 259 Pages). Columbia University Press.

Ronzoni, M., & Valentini, C. (2013). Costruttivismo e Intuizioni Morali: Un Approccio Integrato'. In C. Bagnoli (Ed.), *Che Fare? Nuove Prospettive Filosofiche sull'azione*. Carocci.

Rorty, R. (1989). *Contingency, Irony, and Solidarity*. Cambridge University Press. https://doi.org/10.1017/CBO9780511804397

Rorty, R. (Ed.). (1990). Solidarity or objectivity? In *Objectivity, Relativism, and Truth: Philosophical Papers* (Vol. 1, pp. 21–34). Cambridge University Press. https://doi.org/10.1017/CBO9781139173643.002

Ross, A. (2019). *On Law and Justice* (J. V. H. Holtermann, Ed.; U. Bindreiter, Trans.; 1st ed.). Oxford University Press. https://doi.org/10.1093/oso/9780198716105.001.0001

Ross, W. D. (1931). *The right and the good*. Clarendon Press.

Ross, W. D. (1939). *Foundations of Ethics*. Clarendon Press.

Ruse, M. (1986). *Taking Darwin seriously: A naturalistic approach to philosophy*. Blackwell.

Ryan, A. (1986). *Property and Political Theory* (Reprint edition). Blackwell Pub.

Samuel, J. (2023). Toward a Post-Kantian Constructivism. *Ergo an Open Access Journal of Philosophy*, *9*(0). https://doi.org/10.3998/ergo.3116

Sandel, M. (1982). *Liberalism and the Limits of Justice*. Cambridge University Press.

Sangiovanni, A. (2008). Justice and the Priority of Politics to Morality. *Journal of Political Philosophy*, *16*(2), 137–164. https://doi.org/10.1111/j.1467-9760.2007.00291.x

Sangiovanni, A. (2019). Human Rights in a Kantian Key. *Kantian Review*, *24*(2), 249–261. https://doi.org/10.1017/S1369415419000049

Scanlon, T. M. (1982). Contractualism and utilitarianism. In A. Sen & B. Williams (Eds.), *Utilitarianism and Beyond* (pp. 103–128). Cambridge University Press. https://doi.org/10.1017/CBO9780511611964.007

Scanlon, T. M. (1998). *What we owe to each other*. The Belknap Press of Harvard University Press.

Scanlon, T. M. (2002). Rawls on Justification. In S. Freeman (Ed.), *The Cambridge Companion to Rawls* (pp. 139–167). Cambridge University Press. https://doi.org/10.1017/CCOL0521651670.004

Schachter, O. (1983). Human Dignity as a Normative Concept. *American Journal of International Law*, *77*(4), 848–854. https://doi.org/10.2307/2202536

Schapiro, T. (2021). *Feeling like it: A theory of inclination and will*. Oxford university press.

Schauer, F. (1988). Formalism. *Yale Law Journal*, *97*(4), 509–548.

Schlosberg, D. (1999). *Environmental justice and the new pluralism: The challenge of difference for environmentalism*. Oxford University Press. https://books.google.com/books?hl=en&lr=&id=hXjnCwAAQBAJ&oi=fnd&pg=PR11&dq=info:NTTFa_QVLosJ:scholar.google.com&ots=vQI_HRzYIz&sig=sa_pb3Ot5Q_aNvHtaLJVMW8NAzM

Schneewind, J. B. (1998). *The Invention of Autonomy: A History of Modern Moral Philosophy*. Cambridge University Press.

Schneewind, J. B. (Ed.). (2007). *Moral philosophy from Montaigne to Kant* (Reprinted). Cambridge Univ. Press.

Schroeder, M. (2008). What is the Frege-Geach Problem? *Philosophy Compass*, *3*(4), 703–720. https://doi.org/10.1111/j.1747-9991.2008.00155.x

Sellars, W. (1963). *Science, Perception and Reality*. Humanities Press.

Sen, A. (1984). *Resources, Values and Development*. Basil Blackwell.

Sen, A. (2004). Elements of a Theory of Human Rights. *Philosophy & Public Affairs*, *32*(4), 315–356. https://doi.org/10.1111/j.1088-4963.2004.00017.x

Sensen, O. (2011). *Kant on Human Dignity*. De Gruyter. https://doi.org/10.1515/9783110267167

Sensen, O. (2017). Kant's Constitutivism. In *Realism and Antirealism in Kant's Moral Philosophy: New Essays* (pp. 197–222). De Gruyter. https://doi.org/10.1515/9783110574517-009

Shafer-Landau, R. (1994). Ethical Disagreement, Ethical Objectivism and Moral Indeterminacy. *Philosophy and Phenomenological Research*, *54*(2), 331–344. https://doi.org/10.2307/2108492

Shafer-Landau, R. (2003). *Moral Realism: A Defence*. Oxford University Press. https://doi.org/10.1093/0199259755.001.0001

Shapiro, S. J. (2007). The "Hart-Dworkin" Debate: A Short Guide for the Perplexed. *SSRN Electronic Journal*, 1–54. https://doi.org/10.2139/ssrn.968657

Shapiro, S. J. (2009). What Is the Rule of Recognition (And Does It Exist)? In M. Adler & K. E. Himma (Eds.), *The Rule of Recognition and the U.S. Constitution* (pp. 235–268). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195343298.003.0009

Shiner, R. A. (1992). *Norm and nature: The movements of legal thought*. Clarendon Press.

Shue, H. (1980). *Basic Rights*. Princeton University Press. https://press.princeton.edu/books/paperback/9780691202280/basic-rights

Shue, H. (2020). *Basic Rights*. https://press.princeton.edu/books/paperback/9780691202280/basic-rights

Sidgwick, H. (1874). *The Methods of Ethics* (2011 Ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139136617

Silk, A. (2014). Nietzschean Constructivism: Ethics and Metaethics for All and None. *Inquiry*, *58*, 1–37. https://doi.org/10.1080/0020174X.2013.878666

Simmons, A. J. (2015). Human Rights, Natural Rights, and Human Dignity. In R. Cruft, S. M. Liao, & M. Renzo (Eds.), *Philosophical Foundations of Human Rights* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199688623.003.0007

Simpson, R. M. (2021). Minimalism, Determinacy, and Human Rights. *Canadian Journal of Law & Jurisprudence*, *34*(1), 149–169. https://doi.org/10.1017/cjlj.2020.25

Singer, P. (1976). *Animal Liberation: A New Ethics for Our Treatment of Animals*. Jonathan Cape Ltd. https://www.biblio.com/book/animal-liberation-new-ethics-our-treatment/d/1505186017

Singer, P. (2009). Speciesism and Moral Status. *Metaphilosophy*, *40*(3–4), 567–581. https://doi.org/10.1111/j.1467-9973.2009.01608.x

Singer, P. (2011). *Practical Ethics* (3rd ed.). Cambridge University Press.

Skidmore, J. (2002). Skepticism about Practical Reason: Transcendental Arguments and Their Limits. *Philosophical Studies*, *109*(2), 121–141. https://doi.org/10.1023/A:1016217224647

Skorupski, J. (1998). Rescuing Moral Obligation. *European Journal of Philosophy*, *6*(3), 335–355. https://doi.org/10.1111/1468-0378.00065

Skorupski, J. (2010). *The Domain of Reasons*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199587636.001.0001

Smith, B. (1928). Legal Personality. *The Yale Law Journal*, *37*(3), 283. https://doi.org/10.2307/789740

Smith, M. (1987). The Humean Theory of Motivation. *Mind*, *XCVI*(381), 36–61. https://doi.org/10.1093/mind/XCVI.381.36

Smith, M. (1994). *The moral problem*. Blackwell.

Sreenivasan, G. (2010). Duties and Their Direction. *Ethics*, *120*(3), 465–494. https://doi.org/10.1086/652303

Steinberger, F. (2022). The Normative Status of Logic. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2022). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2022/entries/logic-normative/

Steiner, H. (1994). *An Essay on Rights*. Wiley-Blackwell.

Steiner, H. (2013). Directed Duties and Inalienable Rights. *Ethics*, *123*(2), 230–244. https://doi.org/10.1086/668708

Steiner, H., Kramer, M. H., & Simmonds, N. E. (Eds.). (2002). *A debate over rights: Philosophical enquiries* (Reprinted). Oxford Univ. Press.

Stocker, M. (1976). The Schizophrenia of Modern Ethical Theories. *Journal of Philosophy*, *73*(14), 453–466. https://doi.org/10.2307/2025782

Stone, M. J. (2017). Kant's Apparent Positivism. In S. Kisilevsky & M. J. Stone (Eds.), *Freedom and Force: Essays on Kant's Legal Philosophy*. Bloomsbury Academic.

Street, S. (2008). Constructivism About Reasons. *Oxford Studies in Metaethics*, *3*, 207–245.

Stucki, S. (2023). *One Rights: Human and Animal Rights in the Anthropocene*. Springer International Publishing. https://doi.org/10.1007/978-3-031-19201-2

Sumner, L. W. (1987). *The moral foundation of rights*. Clarendon Press ; Oxford University Press.

Tasioulas, J. (2010). Taking Rights out of Human Rights. *Ethics*, *120*(4), 647–678. https://doi.org/10.1086/653432

Tasioulas, J. (2015). On the Foundations of Human Rights. In R. Cruft, S. M. Liao, & M. Renzo (Eds.), *Philosophical Foundations of Human Rights* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199688623.003.0002

Tenenbaum, S. (2021). *Rational Powers in Action: Instrumental Rationality and Extended Agency*. Oxford University Press USA - OSO.

Tersman, F. (2022). Moral Disagreement. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2022). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2022/entries/disagreement-moral/

Theunissen, L. N. (2020a). *The value of humanity* (First edition). Oxford University Press.

Theunissen, L. N. (2020b). *The value of humanity* (First edition). Oxford University Press.

Thompson, M. (2004). What is it to Wrong Someone?: A Puzzle about Justice. In R. J. Wallace (Ed.), *Reason and value: Themes from the moral philosophy of Joseph Raz*. Clarendon ; Oxford University Press.

Thompson, M. (2008). *Life and action: Elementary structures of practice and practical thought*. Harvard university press.

Thomson, J. J. (1971). A Defense of Abortion. *Philosophy & Public Affairs*, *1*(1), 47–66.

Thomson, J. J. (1990). *The realm of rights*. Harvard Univ. Press.

Timmermann, J. (2015). What's Wrong with "Deontology"? *Proceedings of the Aristotelian Society*, *115*, 75–92.

Timmons, M. (2003). The Limits of Moral Constructivism. *Ratio*, *16*(4), 391–423. https://doi.org/10.1046/j.1467-9329.2003.00230.x

Tomasello, M. (2009). *Cultural Origins of Human Cognition*. Harvard University Press.

Tomassini, F. (2024). Three Models of Natural Right: Baumgarten, Achenwall, and Kant. In C. D. Fugate & J. Hymers (Eds.), *Baumgarten and Kant on the Foundations of Practical Philosophy* (1st ed., pp. 223–240). Oxford University PressOxford. https://doi.org/10.1093/oso/9780192873538.003.0013

Tuck, R. (1982). *Natural Rights Theories: Their Origin and Development*. Cambridge University Press.

Unger, R. M. (2015). *The critical legal studies movement: Another time, a greater task*. Verso.

Valentini, L. (2016). On the Justification of Basic Rights. *Netherlands Journal of Legal Philosophy*, *45*(3), 52–63. https://doi.org/10.5553/njlp/221307132016045003005

Van Jaarsveld, J. (2020). Nussbaum's capability approach and African environmental ethics: Is the African voice heard? *Oxford Development Studies*, *48*(2), 135–147. https://doi.org/10.1080/13600818.2020.1759037

Vatter, M. (2020). Nature's Law or Law's Law? Community of Life, Legal Personhood, and Trusts. In M. de Leeuw & S. van Wichelen (Eds.), *Personhood in the Age of Biolegality: Brave New Law* (pp. 225–245). Springer International Publishing. https://doi.org/10.1007/978-3-030-27848-9_13

Velleman, J. D. (2000a). From Self Psychology to Moral Philosophy. *Philosophical Perspectives*, *14*, 349–377.

Velleman, J. D. (2000b). *The possibility of practical reason* (Reprint.). Oxford Univ. Press.

Velleman, J. D. (2009). *How we get along*. Cambridge University Press. https://doi.org/10.1017/CBO9780511808296

Viafora, C. (1999). Toward a methodology for the ethical analysis of clinical practice. *Medicine, Health Care and Philosophy*, *2*(3), 283–297. https://doi.org/10.1023/A:1009909613820

Villey, M. (2013). *La formation de la pensée juridique moderne* (S. Rials & É. Desmons, Eds.; 2e éd.). Presses universitaires de France.

Waldron, J. (Ed.). (1984). *Theories of rights*. Oxford Univ. Press.

Waldron, J. (1996). Kant's Legal Positivism. *Harvard Law Review*, *109*(7), 1535. https://doi.org/10.2307/1342024

Waldron, J. (2012). *Dignity, rank, and rights*. Oxford University Press.

Waldron, J. (2015). Is Dignity the Foundation of Human Rights? In R. Cruft, S. M. Liao, & M. Renzo (Eds.), *Philosophical Foundations of Human Rights* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199688623.003.0006

Wallace, R. J. (2009). The Publicity of Reasons. *Philosophical Perspectives*, *23*, 471–497.

Waluchow, W. J. (1994). *Inclusive Legal Positivism*. Oxford University Press.

Weber, M. (1989). *Science as vocation* (P. Lassman, I. Velody, & H. Martins, Eds.). Unwin Hyman.

Weinrib, E. J. (2012). *The idea of private law* (First edition). Oxford University Press.

Wellman, C. (1985). *A Theory of rights: Persons under laws, institutions, and morals*. Rowman and Allanheld.

Wellman, C. (1995). *Real Rights*. Oxford University Press.

Wellman, C. (2011a). The Functions of Rights. *ARSP: Archiv Für Rechts- Und Sozialphilosophie / Archives for Philosophy of Law and Social Philosophy*, *97*(2), 169–177.

Wellman, C. (2011b). *The moral dimensions of human rights*. Oxford University Press.

Wenar, L. (2005). The Nature of Rights. *Philosophy & Public Affairs*, *33*(3), 223–252.

Wenar, L. (2013). The Nature of Claim-Rights. *Ethics*, *123*(2), 202–229. https://doi.org/10.1086/668707

Wenar, L. (2023). Rights. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Spring 2023). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2023/entries/rights/

Whitman, J. Q. (2005). 'Human dignity' in Europe and the United States: The social foundations. In G. Nolte (Ed.), *European and US Constitutionalism* (pp. 108–124). Cambridge University Press. https://doi.org/10.1017/CBO9780511493904.008

Wiggins, D. (1987a). *A Sensible Subjectivism*. Oxford: Oxford University Press.

Wiggins, D. (1987b). Deliberation and Practical Reason. In *Needs, Values and Truth: Essays in the Philosophy of Value* (pp. 221–240). Blackwell.

Williams, B. (1979). Internal and External Reasons. In R. Harrison (Ed.), *Rational action: Studies in philosophy and social science* (pp. 101–113). Cambridge University Press.

Williams, B. (1981). *Moral Luck: Philosophical Papers 1973–1980*. Cambridge University Press. https://doi.org/10.1017/CBO9781139165860

Williams, B. (1985). *Ethics and the limits of philosophy*. Harvard University Press.

Wong, D. B. (2008). Constructing Normative Objectivity in Ethics. *Social Philosophy and Policy*, *25*(1), 237–266. https://doi.org/10.1017/S0265052508080096

Wood, A. W. (2014). *The Free Development of Each: Studies on Freedom, Right, and Ethics in Classical German Philosophy*. Oxford University Press.

Wood, A. W., & O'Neill, O. (1998). Kant on Duties Regarding Nonrational Nature. *Aristotelian Society Supplementary Volume*, *72*, 189–228.

Yeomans, C. (2012). *Freedom and Reflection: Hegel and the Logic of Agency*. Oxford University Press.