

An integrative review exploring decision-making processes in forensic psychopathology investigations[☆]

Benedetta Barchielli^{a,*}, Clarissa Cricenti^b, Mario Giantesani^a, Giovanna Parmigiani^c, Stefano Ferracuti^a, Cristina Scarpazza^{d,e}

^a Department of Human Neuroscience, "Sapienza" University of Rome, Italy

^b Department of Psychology, "Sapienza" University of Rome, Italy

^c Department of Law, University of Modena and Reggio Emilia, Italy

^d Department of General Psychology, University of Padova, Italy

^e Translational NeuroImaging and Cognitive Lab, IRCCS S Camillo Hospital, Venice Lido, Venice

ARTICLE INFO

Keywords:

Decision-making
Forensic psychopathology
Bias
Forensic evaluations

ABSTRACT

Forensic psychopathology assessments play a critical role in legal decisions, particularly those related to criminal responsibility. However, despite their influence, the decision-making processes behind these assessments remain largely underexplored and inconsistently supported by empirical evidence. Emerging literature suggests that these decisions are shaped by multiple factors, including individual reasoning styles, team dynamics, and the influence of cognitive biases.

Aim: To explore the factors that influence expert decision-making in forensic psychopathology evaluations, with a focus on cognitive bias, procedural complexity, and the use of structured tools.

Method: An integrative review guided by Cooper's framework was conducted. Eight studies meeting inclusion criteria were analyzed using narrative synthesis and thematic coding.

Results: Three key themes were identified: (1) cognitive bias as a pervasive influence on expert judgment, (2) the iterative and context-dependent nature of forensic decision-making, and (3) inconsistencies between subjective clinical reasoning and structured assessment tools. While structured professional judgment tools and validated instruments showed potential for increasing accuracy and inter-rater reliability, their use remained uneven. Additionally, experienced clinicians often favored subjective judgment over formalized methods, which may limit transparency and reproducibility. The review highlights the need to reframe forensic evaluation as a reflective, evidence-informed process grounded in transparency and standardization. Integrating structured methods and debiasing strategies from other medical disciplines may enhance the validity and fairness of forensic judgments. These findings have implications for clinical training, interdisciplinary collaboration, and legal policy, and underscore the urgency of further research in this underdeveloped area of forensic mental health.

1. Introduction

The decision-making processes underlying forensic psychopathology assessments are complex, variable, and highly case specific. Göranson et al. (2022) have noted that even trained experts can reach divergent conclusions when presented with the same case material. Forensic psychopathology assessments, frequently determinative of legal outcomes, such as criminal responsibility, must be informed by the best available evidence for complex decision-making. Nevertheless, the

empirical evidence base underpinning these processes is limited, particularly concerning how forensic experts make decisions and which factors influence them.

Considering this gap, it is essential to examine the types of information examined by forensic evaluators and consider the extent to which cognitive distortions (systematic errors in judgment) might undermine the objectivity and integrity of legal proceedings. Cognitive distortion, or logical fallacy, refers to the unconscious processes that undermine human thinking in all settings, including judicial

[☆] This article is part of a Special issue entitled: 'DMM&L 2024' published in International Journal of Law and Psychiatry.

* Corresponding author at: Department of Human Neuroscience, Sapienza University of Rome, Viale, Regina Elena, 33400185 Roma, Italy.

E-mail address: benedetta.barchielli@uniroma1.it (B. Barchielli).

applications (Dietrich & Haider, 2015; Dror & Murrle, 2018). Distortion can influence how evidence is collected, interpreted, and assigned weight, possibly contributing to forensic errors.

Prior research in various fields has revealed that such biases often impact expert judgment. Studies in medicine (Einhorn, 1974), psychology (Oskamp, 1965), and law (Ebbesen & Konecni, 1975) have indicated that experts are generally overconfident in their conclusions, without a comparable improvement in accuracy (Shanteau, 1992). For instance, Oskamp (1965) reported that as more information in a clinical case was examined by psychologists, their confidence in their judgments improved, but the accuracy of the conclusions reached an early plateau. Jackson (1986) found that both experts and non-experts demonstrated comparable performance when assessing criminal responsibility and the likelihood of reoffending, highlighting that professional expertise may not necessarily protect against such biases.

Recent studies have provided a comprehensive understanding of this phenomenon. Grøndahl et al. (2009), for instance, reported marked differences in how professionals and laypersons responded to forensic case scenarios. The researchers did find statistically significant differences between laypeople and professionals (psychologists and psychiatrists) in their assessments of forensic cases. Laypeople tended to rate the risk of recidivism, mental illness, and need for treatment higher than experts. The study found that experts assessed forensic cases differently than laypeople, which goes against past research that suggested experts weren't more accurate or useful. These findings challenged previous claims of expert fallibility and led the authors to suggest that such criticism may have been overstated. Göranson et al. (2022) conducted an in-depth examination of how professional groups with differing practice paradigms (forensic psychiatrists, psychologists, and social workers) utilize specific versus general types of information when drawing conclusions regarding the evaluation of severe mental disorders. Although differences emerged in the types of information considered valuable, such as the results from cognitive tests, experts largely relied on a common set of general information sources, which they adapted to the specifics of each case. This finding underscores the complexity and highly individualized nature of forensic decision making.

In their complementary work, Svensson et al. (2022a, 2022b) conducted 38 semi-structured interviews with Swedish forensic psychiatric experts and applied thematic analysis to understand their decision-making process. The results indicate that such processes are not only complex but also iterative, context-dependent, influenced by dynamic interactions among various sources of information, and bounded by situational factors, such as time pressure. Although such intricacy might add validity to conclusions drawn in forensic work, it also makes them vulnerable to cognitive distortions.

These issues are not limited to psychiatric evaluation. Across various forensic science disciplines, including DNA analysis, fingerprint identification, toxicology, and forensic pathology, there is evidence that contextual and cognitive biases can influence expert conclusions (Dror et al., 2021). Dror and Hampikian (2011) demonstrated that even highly objective fields are susceptible to such biases, raising concerns that similar or potentially greater levels of vulnerability may exist within the interpretive domain of forensic psychiatry.

In this context, increasing attention has been directed toward forensic cognitive bias, defined as “the class of effects by which one’s pre-existing beliefs, expectations, motivations, and situational context affect the collection, perception, and interpretation of evidence in criminal cases” (Kassin et al., 2013). This issue is particularly relevant in criminal responsibility assessments, which assist courts in determining whether a defendant should be legally accountable for an offense. Given that such assessments must meet the standard of “beyond reasonable doubt”, it is critical to examine whether their outcomes may be influenced, consciously or unconsciously, by cognitive biases (Scarpazza & Ghidini, 2023).

Evidence also suggests that agreement among forensic evaluators is

often limited. Gowensmith et al. (2012), in their analysis of 216 competency assessments conducted in Hawaii, found that 29 % of cases showed disagreement among three independent evaluators all appointed by the judge. In another sample of 165 sanitary assessments, discrepancies were observed in 45 % of the cases. This inconsistency highlights potential sources of variability in forensic judgments, including differences in expertise, interpretation, case-specific factors, and the possible influence of cognitive bias on evaluators' reasoning.

While other forensic disciplines have made significant progress in identifying bias within their practices, forensic psychiatry remains relatively under researched in this area. Despite the growing acknowledgment of its impact, limited research has systematically examined how cognitive biases influence mental health professionals' judgments in legal proceedings. Moreover, there is uncertainty regarding the extent to which forensic professionals are aware of these biases and whether they actively implement strategies to mitigate their effects (Scarpazza et al., 2021).

This integrative review aims to examine the decision-making processes involved in forensic psychiatric investigations and identify strategies for enhancing the quality of these processes. Specifically, the review seeks to understand the types of information relied upon by experts, how this information and other available resources are utilized, and to delineate the cognitive and procedural mechanisms underlying expert decision-making in this context.

2. Methods

Cooper's (1989) framework was used in this review to direct the integrative methodology. A narrative-descriptive methodology was used, which provided a comprehensive viewpoint and permitted the inclusion of several study designs. Effective integrated reviews inform theory, practice, and policy, in addition to summarizing the state of science (Whittemore & Knafl, 2005).

Cooper's five steps – problem identification, literature search, data evaluation, data analysis, and presentation - ensure a methodical and structured process that produces rigorous, empirically supported, and field-relevant insights.

2.1. Problem identification

In the field of forensic psychopathology, there is still a notable lack of research examining the complex factors that shape how professionals make critical decisions. Literature offers only limited insight into how forensic mental health practitioners, such as psychologists, and psychiatrists, approach and justify such decisions in real-world settings (Neal et al., 2018).

Understanding how these biases can influence or change judgments in clinical and forensic situations is crucial. This knowledge is important for maintaining ethical and professional care standards. It also helps ensure that when force is used, it is done fairly and correctly, especially in situations where a person's freedom and safety might be at risk.

2.2. Literature search stage

To conduct this review, we searched multiple databases, including PubMed, PsycINFO, Scopus, and Web of Science. The search terms focused on key concepts in forensic psychopathology, including investigation, expertise, and decision-making. To ensure thoroughness and accuracy, two researchers independently screened the study titles and abstracts. In cases in which eligibility could not be determined from the abstract, the full text was retrieved and reviewed. Discrepancies in study selection were resolved through discussion with a third reviewer providing input when necessary.

The inclusion criteria targeted a population of psychiatrists, psychologists, and mental health professionals. Eligible studies were required to focus on evaluation of bias, cognitive distortions, and inter-

rater agreement among experts. Inclusion criteria included studies published within the last 20 years. This timeframe was chosen to ensure that findings reflect current diagnostic criteria, legal frameworks, and methodological advancements relevant to contemporary forensic practice. Only studies published in English or that employed either quantitative or qualitative methodologies were considered. Exclusion criteria included studies involving populations outside the field of forensic psychopathology, theses, policy documents, book chapters, commentaries, editorials, and literature reviews. Studies published in languages other than English were also excluded.

Data extraction was performed using a standardized form, enabling the systematic capture of essential study details, including the research focus, participant characteristics, measures used, conclusions, and any identified risk of bias. This process was conducted collaboratively: two reviewers carried out the initial extraction, a third expert resolved any disagreements between the authors during theme extraction through a double-blind process.

In the data evaluation stage, modified Critical Appraisal Skills Programme (CASP; Walsh & Downe, 2006) tools were used to assess quantitative and mixed-method studies. Two reviewers independently appraised each study, focusing on research design, sampling, data collection, ethical considerations, and findings. The appraisal results were compared for consistency, with no major discrepancies noted. Studies receiving grade D, indicating significant methodological flaws

affecting validity and reliability, were excluded.

2.3. Data analysis

Based on Whittemore and Knaf (2005), we employed a constant comparison method. This involved coding the extracted data systematically to identify patterns, themes, and relationships within individual studies and across the dataset.

Fig. 1 outlines the stages of the literature search, study selection, and data analysis, conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Page, McKenzie, et al., 2021; Page, Moher, et al., 2021). Table 1 details the included studies and the prominent themes drawn from each. The findings are indicative of the data analysis and presentation phases of the integrative review, where primary insights are ascertained, documented, and graphically depicted. Eight articles were included in the final synthesis.

Following the extraction of relevant information from the Results and Discussion sections of all included studies, the data was coded systematically. The approach involved an iterative coding process in which patterns were examined both within single articles and across the entire dataset.

Initially, codes were grouped according to the objectives of this review: (1) types of information utilized in forensic psychiatric decision-

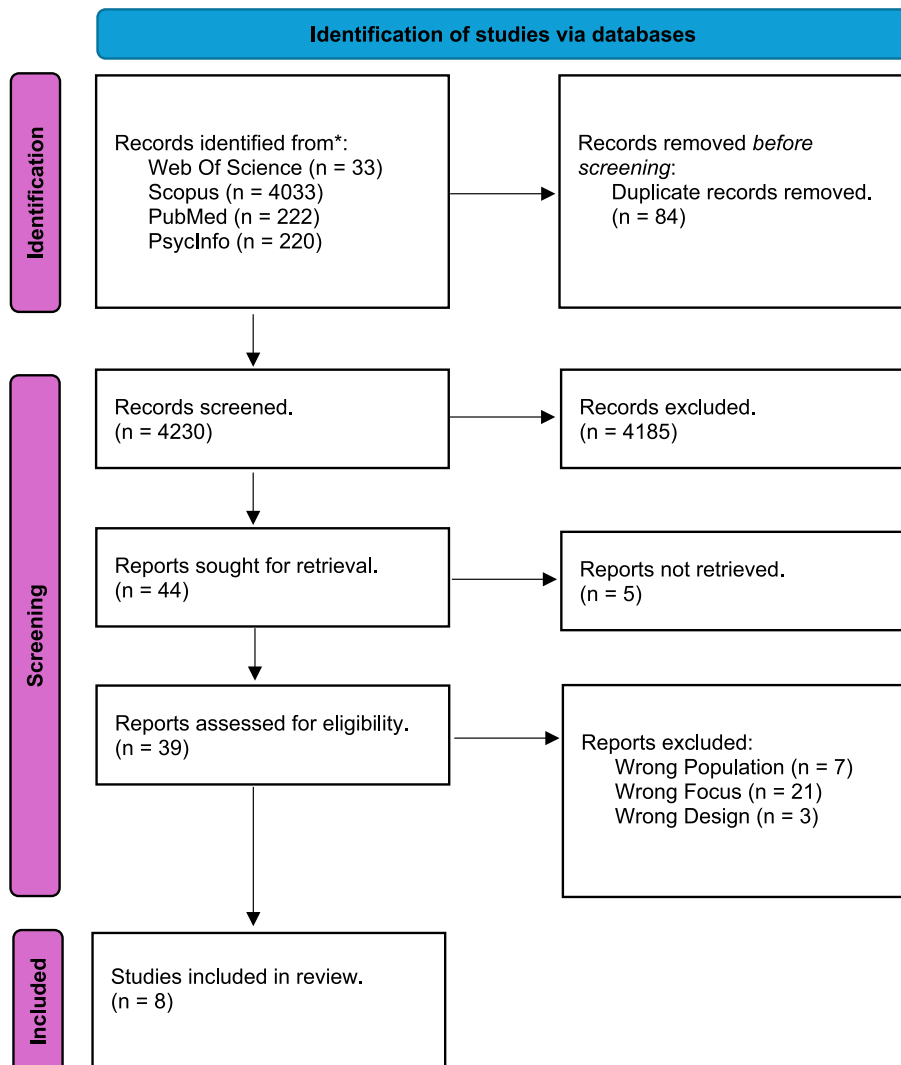


Fig. 1. Flow diagram of integrative review.

Table 1
Summary of studies.

Author & Year	Aim	Setting Participants	Methods	Key Themes	Appraisal Grading
McKee et al. (2007)	Explore hindsight bias among novices vs. forensic experts	104 forensic tribunal hearings	Case vignette analysis	Actuarial vs. clinical judgment, inconsistency in risk assessment application	A
Grøndahl et al. (2009)	Compare judgment accuracy between laypeople and forensic experts	21 psychologists, 14 psychiatrists, 126 laypeople	Case vignette ratings	Expertise effect, lay vs. expert differences, crime severity bias	B
Göranson et al. (2024)	Compare judgment processes between forensic experts and laypersons using psychiatric case vignettes	14 psychiatrists, 21 psychologists, 126 laypersons	Rating of 18 forensic psychiatric case vignettes varying in social, psychiatric, and criminal details	Experts used more information and expressed lower confidence than laypersons; priming effects influenced risk/treatment ratings; experience correlated with lower insanity/treatment ratings and higher risk ratings	B
Svensson et al. (2022a, 2022b)	Explore decision-making process in Swedish forensic psychiatry	38 forensic experts in Sweden	Semi-structured interviews	Team decision-making, cognitive biases, impact of time constraints	A
Göranson et al. (2022)	Study expert information use in forensic psychiatric investigations	41 Swedish forensic professionals	Case vignettes, survey	Core information reliance, case-specific adaptations	A
Weber et al. (2024)	Explore hindsight bias among novices vs. forensic experts	101 participants (novices and experts)	Case vignette analysis	Hindsight bias, effectiveness of awareness-based debiasing	A
Svensson et al. (2024)	Investigate decision-making processes in Swedish forensic psychiatry	27 Swedish forensic experts	Case vignette study	Hypothesis formulation, certainty shifts, diagnostic diversity, absence of confirmation bias	A
Göranson et al. (2024)	Explore diagnostic reasoning and arguments regarding severe mental disorder	41 forensic psychiatrists, psychologists, social workers in Sweden	Qualitative vignette study	Dual process theory, illness scripts, psychotic functioning as core SMD basis, case-specific adaptations	A

making, (2) integration and application processes, and (3) decision quality enhancement strategies. Through a comparison of the iterations, these categories were reduced to three broad themes: (i) influence of cognitive biases in decision-making processes; (ii) complexity and iterative nature of decision-making, (iii) discrepancies between clinical judgment and objective indicators.

2.4. Theme

1: Influence of Cognitive Biases in Decision-Making Processes.

Cognitive bias as a pervasive element in forensic expert decision-making emerged as a consistent theme across multiple studies. Despite their professional training and experience, forensic experts remain vulnerable to systematic distortions in judgment. Svensson et al. (2022a, 2022b), for example, emphasized that decision-making processes are sensitive to contextual pressures, particularly time constraints. They found that the sequence in which information is presented—known as priming—can significantly shape judgments of dangerousness and criminal responsibility. A lack of information was also found to decrease experts' confidence in their evaluations, which contrasts with the overconfidence typically exhibited by non-experts, and suggests a more cautious, yet still fallible, cognitive approach among professionals. In a follow-up study using forensic case vignettes, Svensson et al. (2024) explored these cognitive dynamics further. They didn't see clear evidence of confirmation bias, meaning they didn't observe experts actively interpreting new information in a way that supported their initial beliefs. However, they did notice that the experts held on to their first diagnosis even after new information was introduced. This behavior doesn't necessarily confirm confirmation bias, but it could suggest a degree of cognitive rigidity (an unwillingness to revise conclusions once they've been made). So, while the experts didn't distort or ignore the new information, they did tend to stick with their initial impressions. This may indicate a subtle, implicit form of bias, even if it wasn't overtly expressed. Weber et al. (2024) compared experts and novices in their susceptibility to hindsight bias, the distortion of judgment based on known outcomes. Although experts were less prone to this bias than non-professionals, they still showed measurable benefit from debiasing interventions, reinforcing the idea that expertise alone does not confer immunity to cognitive distortions. As one expert reflected, "being aware of the bias isn't always enough; sometimes it creeps in regardless"

(Weber et al., 2024, p. 13). In a related observation, McKee et al. (2007) examined the use of structured risk assessment tools and found that although experts claimed to rely on actuarial indicators, their recommendations frequently diverged from the numerical outcomes provided. This inconsistency points to a disjunction between stated decision criteria and actual practice, possibly reflecting the influence of implicit biases or retrospective rationalization.

2.5. Theme

2: Complexity and Iterative Nature of Decision-Making.

The complexity and iterative nature of decision-making was another recurring theme across the reviewed literature. Rather than following a linear or protocol-driven pathway, forensic decision-making was consistently portrayed as dynamic, fluid, and deeply context-dependent. Svensson et al. (2022a, 2022b) described the decision-making process as largely unstandardized and shaped by the internal culture and working methods of each forensic unit. Evaluations were approached as unique events, and conclusions emerged through ongoing negotiation among members of multidisciplinary teams, including psychiatrists, psychologists, and social workers. The study emphasized that collaboration, while enriching, also introduced variability, particularly in how team members weighed and interpreted evidence. Building on this, Svensson et al. (2024) examined expert reasoning using standardized vignette methodology. They found that even under controlled conditions, professionals formulated, revised, and refined diagnostic hypotheses in a cyclical manner. Despite shared training and similar tasks, expert judgments varied significantly depending on individual reasoning styles. This finding pointed both to a cognitive flexibility in navigating complex cases, and to a broader absence of shared, uniform decision-making protocols. Göranson et al. (2022) echoed these observations by emphasizing the need for a structured integration of multiple sources of information (eg. interviews, clinical files, and behavioral observations) in forensic assessments. They applied the conceptual framework proposed by Dror and Murrie (2018), the Hierarchy of Expert Performance (HEP), which outlines ascending levels of expertise in the selection, weighting, and synthesis of data. This model supports the view that decision-making evolves through reflective practice and deliberate engagement with diverse forms of evidence. Further elaborating this view, Göranson et al. (2024) described forensic assessment as a

narrative and interpretive process, in which experts navigate competing explanations to determine whether a Severe Mental Disorder (SMD) contributed to the criminal behavior. The authors noted that the construction of these narratives is shaped by the professional background of the assessor, psychiatrist, psychologist, or social worker, adding another layer of variability to the process. Similarly, Grøndahl et al. (2009), using clinical vignettes, found that forensic experts engaged with information in more complex ways than laypeople. Their study revealed that expert evaluations were not merely more informed, but also differed in how key elements such as psychiatric history or crime severity were prioritized. Mandarelli et al. (2019) also contributed to this discussion by analyzing how sociodemographic, psychopathological, and criminological variables influence assessments of criminal responsibility and dangerousness. Their findings revealed that experts' conclusions were often shaped by a mixture of clinical factors and case-specific nuances. Importantly, the study highlighted the presence of a "time bias", a discrepancy between the mental state at the time of the offense and at the time of the evaluation, which further complicates the process. This temporal complexity reinforces the iterative and interpretive nature of forensic decision-making, where conclusions emerge not from fixed criteria but through the reconstruction of past mental states based on incomplete, retrospective data.

2.6. Theme

3: Discrepancies Between Clinical Judgment and Objective Indicators.

The discrepancy between clinical judgment and structured indicators emerged as a salient theme across multiple studies. Despite the widespread availability of standardized tools designed to enhance consistency and objectivity, forensic practitioners frequently diverged from these instruments in favor of more intuitive or narrative-based approaches. McKee et al. (2007) provided a clear example of this phenomenon, observing that although clinicians routinely reported using actuarial tools such as the Violence Risk Appraisal Guide (VRAG), their final recommendations to forensic tribunals often contradicted the scores those tools generated. This disconnect revealed a preference for subjective interpretation, in which expert narratives were privileged over algorithmic outputs. The authors suggested that such deviations were rarely acknowledged explicitly, pointing to a potential overestimation of the objectivity clinicians attribute to their own decision-making. Grøndahl et al. (2009) supported this view through a vignette-based study comparing forensic psychologists, psychiatrists, and laypersons. They found that not only did experts assign different weights to key case components than non-experts, but also that substantial variability existed among professionals themselves. This lack of convergence, even within expert groups, cast doubt on the assumed consistency of professional judgment and highlighted the limited reliance on shared, structured frameworks. As the authors noted, "experience does not appear to standardize reasoning, but rather diversifies it." Weber et al. (2024) further expanded on this tension by exploring how professional background and cognitive style influence the use of structured tools. His findings indicated that more experienced assessors were less likely to employ formalized instruments, instead relying on intuitive pattern recognition. Interestingly, the study also demonstrated that even among trained experts, exposure to debiasing interventions increased receptivity to structured guidance, suggesting that the divergence from objectivity is not fixed but modifiable. One participant explained, "I've been doing this long enough to trust my instincts, but the training made me realize how often those instincts align with bias" (Weber et al., 2024, p. 17). Together, these findings underscore an enduring tension within forensic assessment: while structured tools are designed to mitigate variability and enhance reliability, clinical intuition, often shaped by experience and case narrative, continues to dominate the decision-making landscape.

3. Discussion

Overall, the emerging themes from this review suggest that forensic psychopathology experts play a crucial role in the judicial system by providing insights that can significantly influence legal outcomes. Their decision-making processes are complex and involve evaluating various types of evidence under conditions of uncertainty, as illustrated in Table 1. Importantly, these processes are not only shaped by individual expertise but also by the systemic and structural frameworks in which evaluations take place. Differences in how national forensic systems are organized may therefore have a substantial impact on both the consistency and the quality of expert assessments. Pouls et al. (2022), in a recent review, compare how different countries conduct criminal responsibility assessments. Among the jurisdictions examined are Canada, England and Wales, the Netherlands, and Sweden, all of which appear to have well-structured and clearly defined systems for managing such evaluations. The Swedish system stands out for its rigor: all assessments are conducted by a governmental agency employing its own internal staff. It is the court that formally commissions the evaluation from the agency, rather than from individual experts. At the opposite end of the spectrum, the situation in France appears considerably less organized, lacking clear criteria for the selection of experts or for the format of expert reports. Similarly, in Turkey, forensic psychiatry is not officially recognized as a subspecialty of psychiatry; nevertheless, it constitutes a substantial part of psychiatric practice in the country, while still lacking structured guidelines in this field (Balcioglu et al., 2024). We decided to provide an overview of different countries, since the studies included in the review are mainly situated in the Swedish context, and this should be considered when interpreting the findings. Differences in quality standards across countries may directly influence how assessments are conducted and, consequently, the decisions that stem from them (Pouls et al., 2022). Therefore, findings primarily reflect Nordic contexts; applicability to adversarial systems requires verification.

A first critical issue identified in the literature is the influence of cognitive biases on expert decisions. Forensic decision-making is susceptible to biases and reliability issues. Experts are not entirely reliable or unbiased, as their decisions can be influenced by extraneous context and the order in which information is presented (Dror & Kukucka, 2021; Otgaar et al., 2024). These biases can negatively impact the reliability and validity of forensic evaluations. Time constraints and other contextual factors can also reduce the quality of decisions (Raharjanti et al., 2021; Svensson et al., 2022a, 2022b). Recent findings by Svensson et al. (2024) highlight the role of cognitive rigidity rather than overt confirmation bias. While confirmation bias reflects the selective search for or interpretation of evidence supporting prior beliefs, cognitive rigidity refers to a reluctance to update initial hypotheses, even in the presence of contradictory evidence. This subtle but critical distinction points to a deeper mechanism of evaluative bias, linked to difficulties in cognitive flexibility, that deserves more attention in forensic contexts. An effective strategy for mitigating the influence of cognitive distortions involves the deliberate and systematic consideration of alternative diagnostic hypotheses. Despite its potential, this approach remains infrequently applied in forensic contexts, where evaluators rarely articulate the rationale for excluding specific diagnostic possibilities. However, research indicates that explicitly addressing alternative explanations can play a critical role in reducing susceptibility to bias, thereby enhancing the objectivity and reliability of expert judgment (Hirt & Markman, 1995). Nevertheless, strategies developed in other areas of medicine, particularly debiasing techniques, offer valuable models that can be adapted to the forensic context. Debiasing in forensic psychological assessment requires structured strategies beyond self-awareness, including transparent documentation of information sources and sequence, formulation of alternative hypotheses, and the use of standardized tools. Masking irrelevant contextual information through methods such as Linear Sequential Unmasking-Expanded (LSU-E) reduces cognitive contamination, while structured reasoning

frameworks (e.g., Claim–Evidence–Reasoning) enhance transparency and replicability (Oberlader et al., 2025). These interventions, which aim to reduce cognitive distortions through structured reasoning and reflective processes, have shown promise in improving decision quality in high-stakes evaluative settings such as forensic psychopathology (Raharjanti et al., 2021). Biases in forensic expert decision-making have been an interest, with several evaluations highlighting the need to address these cognitive and human factors. As Neal et al. (2018), much of what we know about cognitive and social-cognitive biases in expert decision-making comes from fields other than forensic mental health. Although interest in this area is increasing, many studies focused on forensic contexts are still preliminary and remain unpublished or have yet to undergo peer review. In contrast, more robust evidence of these biases has been established in areas such as forensic science, healthcare, and legal decision-making. This gap underscores the need for rigorous, ecologically valid research specifically within forensic mental health to better understand how such biases may shape expert evaluations (Vredeveltdt et al., 2024). It's also worth considering whether the nature of forensic work, with its ethical demands and high consequences, encourages greater objectivity, potentially reducing susceptibility to bias. On the other hand, the adversarial nature of legal proceedings might amplify certain cognitive distortions, making experts in this field particularly vulnerable to the so-called allegiance effect (Neal, 2016). However, these results were not highlighted in the studies included in this review and therefore would require further empirical validation.

A second theme fundamental to understanding expert evaluations concerns the complexity of forensic decision-making. Forensic psychopathology experts often work in multi-professional teams, where they are required to formulate and revise diagnostic hypotheses based on new information. This iterative process involves significant variation in the hypotheses considered pertinent by different experts, which can lead to diverse opinions throughout and at the end of investigations (Svensson et al., 2022a, 2022b; Svensson et al., 2024). Experts consider a broad range of psychiatric diagnoses and adapt their reasoning to case-specific characteristics, affecting the hypotheses they generate and the information they require (Göranson et al., 2024). For instance, psychosis is often considered a hypothesis across various cases, while other diagnoses are case-specific. Additionally, sociodemographic, psychopathological, and criminological characteristics can affect decisions on criminal responsibility and social dangerousness, with more severe psychiatric symptoms often leading to judgments regarding social dangerousness (Mandarelli et al., 2019).

An additional consideration pertains to the observed inclination among seasoned clinicians to prioritize subjective judgment over the consistent application of standardized assessment tools. Nonetheless, empirical evidence suggests that the use of structured psychodiagnostics tools can significantly enhance inter-rater reliability, even among professionals with extensive clinical expertise (Miller et al., 2015). Within this framework, scientific rigor in forensic evaluation is not achieved through adherence to a single method, but rather through the development of a coherent interpretive hypothesis—one that integrates all available evidence without resorting to arbitrary exclusions. Although clinicians' judgments have been shown to align with patient self-reports, particularly when information is clear, discrepancies often emerge in the presence of ambiguity, leading to a tendency toward conservative evaluations (Defife et al., 2010). As highlighted in the literature, the use of structured tools is widespread across forensic contexts; however, there remains significant heterogeneity, with over 286 different tools reported (Neal et al., 2018), and a lack of standardization that limits comparability across assessments. Structured approaches in forensic psychiatry primarily rely on structured risk assessment instruments (SRAIs), which outperform unstructured clinical judgment in forecasting violent and sexual recidivism (Kamorowski et al., 2022). Widely used tools include Structured professional judgment (SPJ) instruments (e.g., HCR-20 V3, HKT-30/HKT-R, SAPROF, SVR-20) and actuarial instruments (e.g., Static-99/99R, often complemented by STABLE-2007

and ACUTE-2007). In practice, evaluators also employ psychopathy measures (e.g., PCL-R/PCL:SV) as adjuncts to SRAIs to refine risk formulation and treatment planning (Kamorowski et al., 2022). SPJ instruments and validated symptom validity tests, in contrast, have been shown to enhance diagnostic accuracy while reducing reliance on subjective clinical impression (Ng et al., 2021). In this regard, the adoption of structured decision-making tools, along with clearly defined criteria, has the potential to significantly improve the consistency and transparency of forensic evaluations (Weber et al., 2024).

One of the main advantages of an integrative review is its ability to draw on multiple research methods, allowing for a comprehensive and nuanced understanding of the evidence. This approach offers both depth and breadth without prioritizing evidence based on strict methodological hierarchies. However, the inclusion of varied methodologies can also be seen as a drawback, potentially affecting consistency, rigor, and increasing the possibility of bias. Although only peer-reviewed studies were included, the review employed a broad search strategy to capture all significant literature. All authors participated in the data analysis and theme identification process until full agreement was achieved. An additional strength of this review is the consistency of its findings with existing literature, which supports the credibility of the methodology and confirms key factors influencing mental health experts' decisions regarding forensic decisions. The tension between expert intuition and structured tools may, at least in part, be explained by the well-documented phenomenon of overconfidence bias, whereby individuals display greater confidence in their judgments than is warranted by their actual accuracy (Rath et al., 2025). This illusion of validity, pervasive even among trained medical professionals, can lead forensic experts to overvalue their personal experience while underestimating the utility of structured instruments, sometimes perceived as too rigid to capture case-specific complexities, a similar situation can be observed not only in forensic expert decision-making but also in personnel selection processes (Kausel et al., 2016).

Future studies should focus on balancing the implementation of bias-reduction strategies with practical considerations. Overly stringent measures may complicate workflows or undermine effective existing practices, whereas proportionate strategies can reduce contextual bias without creating inefficiencies. Research should therefore prioritize evaluating the cost–benefit ratio and workflow implications of such interventions, ensuring that measures are both effective and sustainable across forensic contexts. In addition, the studies reviewed underscore that assessment is widely regarded as the cornerstone of effective forensic practice, demanding the systematic collection and integration of relevant information to guide judgments regarding intervention and management. Given the profound implications that forensic evaluations may hold for individuals' lives, clinicians carry an ethical obligation to ensure that their inquiries are both meaningful and empirically grounded. Future research should place greater emphasis on examining how cultural and contextual factors, such as ethnicity, traditions, language, and religious values, shape both evaluators' judgments and evaluatees' presentations, ideally through cross-national comparative studies.

3.1. Strengths and limitations

One of the main advantages of an integrative review is its ability to draw on multiple research methods, allowing for a comprehensive and nuanced understanding of the evidence. This approach offers both depth and breadth without prioritizing evidence based on strict methodological hierarchies. However, the inclusion of varied methodologies can also be seen as a drawback, potentially affecting consistency, rigor, and increasing the possibility of bias. The heterogeneity of the included studies represents, also, an intrinsic limitation of the methodology adopted. Since many of the studies employed different experimental designs, it was not possible to conduct a quantitative synthesis such as a meta-analysis, which in turn prevents the estimation of the statistical magnitude of bias effects across studies. Although only peer-reviewed

studies were included, the review employed a broad search strategy to capture all significant literature. Moreover, most of the available evidence originates from Northern European contexts, highlighting the need for further applied research in countries with different legal traditions and systemic frameworks. Consequently, the findings of this review cannot be considered fully generalizable across diverse legal and cultural settings. All authors participated in the data analysis and theme identification process until full agreement was achieved. An additional strength of this review is the consistency of its findings with existing literature, which supports the credibility of the methodology and confirms key factors influencing mental health experts' decisions regarding forensic decisions.

3.2. Implications for practice

Decision-making in forensic psychopathology carries significant legal and personal consequences. However, current practices are often affected by inconsistencies and bias. Addressing these limitations is necessary to support more reliable, valid, and ethically sound evaluations. Forensic experts' decisions are shaped by a combination of clinical reasoning, contextual pressures, and the way information is presented. These elements can introduce variability and reduce fairness, especially when subjective judgment overrides standardized approaches. Experts may interpret the same information differently, leading to divergent conclusions and reduced inter-rater reliability. In some cases, professionals may also be unfamiliar with updated, dimensional models of psychopathology, which limits the depth and objectivity of evaluations. Several strategies have been identified to reduce these risks and improve fairness. Structured and standardized procedures, such as testing multiple diagnostic hypotheses, updating conclusions based on new data, and using defined decision frameworks, can improve consistency and limit errors. Techniques designed to reduce bias, including limiting access to irrelevant information, controlling the sequence in which evidence is reviewed, and using independent raters, have also shown positive effects. In addition, adopting dimensional models of psychopathology can provide a more nuanced understanding of clinical profiles. Forensic psychopathology experts' decision-making is shaped by interrelated factors, including the complexity of clinical judgments and the influence of cognitive bias. While experts are expected to provide objective evaluations, they are often unaware of how unconscious processes and situational pressures may affect their reasoning. This highlights the importance of targeted training programmes that increase awareness of the challenges inherent in forensic decision-making and promote reflective, evidence-based practices.

CRedit authorship contribution statement

Benedetta Barchielli: Writing – original draft, Methodology. **Clarissa Cricenti:** Writing – original draft, Methodology, Conceptualization. **Mario Giancesani:** Writing – review & editing, Investigation. **Giovanna Parmigiani:** Writing – review & editing, Supervision. **Stefano Ferracuti:** Writing – review & editing. **Cristina Scarpazza:** Writing – review & editing, Supervision, Conceptualization.

Funding

This study received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Declaration of competing interest

None.

References

- Balcioglu, Y. H., Oncu, F., & Kennedy, H. G. (2024). Forensic psychiatry in Türkiye. *International Review of Psychiatry*, 36(7), 739–755.
- Cooper, H. M. (1989). *Integrating research: A guide for literature reviews*. Sage Publications, Inc.
- Defife, J., Drill, R., Nakash, O., & Westen, D. (2010). Agreement between clinician and patient ratings of adaptive functioning and developmental history. *The American Journal of Psychiatry*, 167(12), 1472–1478. <https://doi.org/10.1176/appi.ajp.2010.09101489>
- Dietrich, A., & Haider, H. (2015). Human creativity, evolutionary algorithms, and predictive representations: The mechanics of thought trials. *Psychonomic Bulletin & Review*, 22, 897–915.
- Dror, I., & Kukucka, J. (2021). Linear Sequential Unmasking–Expanded (LSU-E): A general approach for improving decision making as well as minimizing noise and bias. *Forensic Science International: Synergy*, 3. <https://doi.org/10.1016/j.fs SYN.2021.100161>
- Dror, I., Melinek, J., Arden, J. L., Kukucka, J., Hawkins, S., Carter, J., & Atherton, D. S. (2021). Cognitive bias in forensic pathology decisions. *Journal of Forensic Sciences*, 66(5), 1751–1757.
- Dror, I., & Murrle, D. (2018). A hierarchy of expert performance applied to forensic psychological assessments. *Psychology, Public Policy, and Law*, 24, 11–23. <https://doi.org/10.1037/law0000140>
- Dror, I. E., & Hampikian, G. (2011). Subjectivity and bias in forensic DNA mixture interpretation. *Science & Justice*, 51(4), 204–208.
- Ebbesen, E. B., & Konecni, V. J. (1975). Decision making and information integration in the courts: The setting of bail. *Journal of Personality and Social Psychology*, 32(5), 805.
- Einhorn, H. J. (1974). Expert judgment: Some necessary conditions and an example. *Journal of Applied Psychology*, 59(5), 562.
- Göranson, L., Svensson, O., Andiné, P., Bromander, S., Ask, K., Bagge, A., & Karlén, H. (2024). Which diagnoses and arguments regarding severe mental disorder do forensic psychiatric experts in Sweden consider in different cases? A qualitative vignette study. *International Journal of Law and Psychiatry*, 96, Article 102003. <https://doi.org/10.1016/j.ijlp.2024.102003>
- Göranson, L., Svensson, O., Andiné, P., Bromander, S., Bagge, A. S. L., & Karlén, M. H. (2022). Decision-making within forensic psychiatric investigations: The use of various information sources by different expert groups to reach conclusions on legal insanity. *Frontiers in Psychiatry*, 13, Article 822519.
- Gowensmith, W. N., Murrle, D. C., & Boccaccini, M. T. (2012). Field reliability of competence to stand trial opinions: How often do evaluators agree, and what do judges decide when evaluators disagree? *Law and Human Behavior*, 36(2), 130.
- Grøndahl, P., Grønnerød, C., & Sexton, J. (2009). A comparative case vignette study of decision making in forensic psychiatric cases. *International Journal of Forensic Mental Health*, 8(4), 263–270.
- Hirt, E. R., & Markman, K. D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology*, 69(6), 1069.
- Jackson, M. W. (1986). Psychiatric decision-making for the courts: Judges, psychiatrists, lay people? *International Journal of Law and Psychiatry*, 9(4), 507–520.
- Kamorowski, J., de Ruitter, C., Schreuder, M., Ask, K., & Jelčić, M. (2022). Forensic mental health practitioners' use of structured risk assessment instruments, views about bias in risk evaluations, and strategies to counteract it. *International Journal of Forensic Mental Health*, 21(1), 1–19.
- Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, 2(1), 42–52.
- Kausel, E. E., Culbertson, S. S., & Madrid, H. P. (2016). Overconfidence in personnel selection: When and why unstructured interview information can hurt hiring decisions. *Organizational Behavior and Human Decision Processes*, 137, 27–44.
- Mandarelli, G., Carabellese, F., Felthous, A., Parmigiani, G., Del Casale, A., Catanesi, R., ... Ferracuti, S. (2019). The factors associated with forensic psychiatrists' decisions in criminal responsibility and social dangerousness evaluations. *International Journal of Law and Psychiatry*, 66, Article 101503. <https://doi.org/10.1016/j.ijlp.2019.101503>
- McKee, S. A., Harris, G. T., & Rice, M. E. (2007). Improving forensic tribunal decisions: The role of the clinician. *Behavioral Sciences & the Law*, 25(4), 483–495. <https://doi.org/10.1002/bsl.768>
- Miller, D. J., Spengler, E. S., & Spengler, P. M. (2015). A meta-analysis of confidence and judgment accuracy in clinical decision making. *Journal of Counseling Psychology*, 62(4), 553.
- Neal, T. M. (2016). Are forensic experts already biased before adversarial legal parties hire them? *PLoS one*, 11(4), Article e0154434.
- Neal, T. M., Hight, M., Howatt, B. C., & Hamza, C. (2018). The cognitive and social psychological bases of bias in forensic mental health judgments. *Advances in Psychology and Law*, 3, 151–175.
- Ng, W., Mattos, L., Coffey, C., Molina, S., Gottfried, E., & Glassmire, D. (2021). The association between clinicians' initial judgments of feigning and outcomes on symptom validity measures among pretrial forensic psychiatric inpatients. *International Journal of Law and Psychiatry*, 76, Article 101698. <https://doi.org/10.1016/j.ijlp.2021.101698>
- Oberlader, V., Quinten, L., Schmidt, A. F., & Banse, R. (2025). How can I reduce bias in my work? Discussing debiasing strategies for forensic psychological assessments. *Professional Psychology: Research and Practice*, 56(3), 211–221. <https://doi.org/10.1037/pro0000615>
- Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology*, 29(3), 261.

- Otgaard, H., De Beuf, T. L., Sauerland, M., & Schincariol, A. (2024). Evaluating the validity of testimony: The role of the order of evidence. *Forensic Science International: Synergy*, 9, Article 100562.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *bmj*, 372.
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *bmj*, 372.
- Pouls, C., Jeandarme, I., Al-Taïar, H., Bradford, J., Canton, W., Kristiansson, M., ... Konrad, N. (2022). Criminal responsibility evaluations: Benchmarking in different countries. *International Journal of Law and Psychiatry*, 81, Article 101775.
- Raharjanti, N. W., Wiguna, T., Purwadianto, A., Soemantri, D., Bardosono, S., Poerwandari, E. K., ... Levania, M. K. (2021). Clinical reasoning in forensic psychiatry: Concepts, processes, and pitfalls. *Frontiers in Psychiatry*, 12, Article 691377.
- Rath, H. J., Rocha, B., Smith, A. M., & Smalarz, L. (2025). Does the forensic filler-control method reduce examiner overconfidence? an experimental investigation using mock fingerprint examiners. *Behav. Sci.*, 15(9), 1191.
- Scarpazza, C., & Ghidini, G. (2023). I bias cognitivi e la loro influenza sull'esito processuale. *Giornale Italiano di Psicologia*, 50(4), 757–780.
- Scarpazza, C., Zampieri, I., Miolla, A., Melis, G., Pietrini, P., & Sartori, G. (2021). A multidisciplinary approach to insanity assessment as a way to reduce cognitive biases. *Forensic Science International*, 319, Article 110652. <https://doi.org/10.1016/j.forsciint.2020.110652>
- Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes*, 53(2), 252–266.
- Svensson, O., Andiné, P., Bromander, S., Ask, K., Bagge, A. S. L., & Karlén, M. H. (2022a). The decision-making process in Swedish forensic psychiatric investigations. *International Journal of Law and Psychiatry*, 80, Article 101709.
- Svensson, O., Andiné, P., Bromander, S., Ask, K., Bagge, L., & Karlén, H. (2022b). The decision-making process in Swedish forensic psychiatric investigations. *International Journal of Law and Psychiatry*, 80, Article 101709. <https://doi.org/10.1016/j.ijlp.2021.101709>
- Svensson, O., Andiné, P., Bromander, S., Ask, K., Bagge, L., & Karlén, H. (2024). Experts' decision-making processes in Swedish forensic psychiatric investigations: A case vignette study. *International Journal of Law and Psychiatry*, 92, Article 101947. <https://doi.org/10.1016/j.ijlp.2023.101947>
- Vredevelde, A., van Rosmalen, E. A., Van Koppen, P. J., Dror, I. E., & Otgaard, H. (2024). Legal psychologists as experts: Guidelines for minimizing bias. *Psychology, Crime & Law*, 30(7), 705–729.
- Walsh, D., & Downe, S. (2006). Appraising the quality of qualitative research. *Midwifery*, 22(2), 108–119.
- Weber, M. A., Johnson, R. D., & Thompson, L. P. (2024). Hindsight bias in forensic mental health novices and experts. *Journal of Forensic Psychology Practice*, 24(1), 1–20. <https://doi.org/10.1080/24732850.2024.2396991>
- Whittemore, R., & Knaf, K. A. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, 52, 546–553.