Pain and fear in the eyes: gaze dynamics predicts social anxiety from fear generalisation

Sabrina Patania[1 [0000−0001−7279−723X]], Alessandro D'Amelio[1 [0000−0002−8210−4457]], Vittorio Cuculo[2 [0000−0002−8479−9950]], Matteo Limoncini[1 [0009−0008−2649−8798]], Marco Ghezzi[1 [0009−0000−0909−7900]], Vincenzo Conversano[1 [0000−0002−2139−5166]], and Giuseppe Boccignone[1 [0000−0002−5572−0924]]

[1] PHuSe Lab - Università degli Studi di Milano
{sabrina.patania, alessandro.damelio, giuseppe.boccignone}@unimi.it
{matteo.limoncini, marco.ghezzi3, vincenzo.conversano}@studenti.unimi.it
[2] AImageLab - Università degli Studi di Modena e Reggio Emilia
{vittorio.cuculo}@unimore.it

Abstract. This study presents a systematic approach for analyzing eye movements in the context of fear generalisation and predicting Social Interaction Anxiety Scale (SIAS) scores. Leveraging principles from foraging theory, we introduce a composite Ornstein-Uhlenbeck (O-U) process as a computational model for social anxiety assessment based on eye-tracking data. Through Bayesian analysis, we infer the model parameters and identify a feature set for SIAS score prediction. The results demonstrate the effectiveness of our approach, achieving promising performance using Random Forest (RF) classification. This research offers a novel perspective on gaze analysis for social anxiety assessment and highlights the potential of gaze behaviour as a valuable modality for psychological evaluation.

Keywords: Fear generalisation · Pain · Eye movements · Visual foraging · Anxiety disorders.

1 Introduction

If you have been painfully bitten by a vicious dog, you might have acquired fear of all dogs therefrom conceptualised as those harmful barking animals with sharp teeth and claws, four legs and a tail. Under such circumstances, regrettably, you are contending with a fear (over)generalisation problem.

Fear generalisation (FG) describes the phenomenon that learnt fear is not restricted to those exact stimuli with which an aversive experience was originally paired (the specific wicked black beast) but it spreads to perceptually or conceptually similar ones (all black dogs or even all dogs, see Fig. 1) [13].

Clearly, the ability to learn which stimuli in the environment signal threat has an important adaptive advantage (to initiate appropriate defensive responses); a hallmark of human cognition is indeed the adeptness to extract conceptual

Fig. 1. Fear learning and generalisation. An aversive episode pairs the perceptual stimulus (conditioned stimulus CS) of a black dog with a painful bite (unconditioned stimulus, US). Learnt fear can subsequently spread over a gradient of harmless stimuli (generalisation stimuli, GSs) more or less similar to the original CS

knowledge from a learning episode [33]. Yet, excessive generalisation may become maladaptive and pathological. Crucially, the overgeneralisation of fear to harmless stimuli or contexts might even turn into a burden to daily life and characteristic of several anxiety disorders [12].

Fear learning and its generalisation effects are usually investigated using the fear-conditioning paradigm [13,16]. Learning is fostered by pairing a perceptual stimulus with an aversive one (e.g., an electric shock) and then the extent of subject's generalisation to stimuli perceptually similar to the original one is assessed via subject's shock expectancy and behavioural/physiological measures (see Section 2). To such end, extensive research has been conducted by using geometric shape stimuli (circles, triangles, etc). Fear learning *in vivo*, however, hardly involves such simple sensory cues; thus, in order to increase the ecological validity of FG studies, recent works [1,30,29] have considered faces as suitable targets. Notably, Reutter and Gamer [29] have provided experimental evidence that the extent of explicit fear generalisation is related to individual patterns of attentional deployment. Precisely, by using visual facial stimuli, eye-tracked participants who dwelled on the distinguishing facial features faster and for longer periods of time were likely to exhibit less fear generalisation. Their analyses though were based on classic measurements (latency of the first fixation, dwell time, etc.); gaze dynamics of viewing behaviour was only indirectly considered.

In this note, based on Reutter and Gamer's publicly available experimental data, we propose a model-based Bayesian analysis of participants' attention deployment (cfr. Section 3). Most important, we take a step further by exploiting the inferred model parameters to predict each participant's social anxiety level, by relying on data gathered in the same experiment via the Social Interaction Anxiety Scale (SIAS) questionnaire [29]. Results of our analysis are reported in Section 4; altogether, it is shown that gaze dynamics is effective at predicting individual's social anxiety in the fear conditioning context. To the best of our knowledge, the approach we present is novel in this direction.

Beyond the theoretical appeal of the problem, the motivation behind our work stems from the very fact that overgeneralisation of fear behaviours is common in many mental health disorders, including specific phobias, obsessive-compulsive disorder, panic disorder, generalised anxiety disorder, and post-traumatic stress disorder (PTSD) [13]. Further, in the chronic-pain condition, spreading of fear to safe movements may lead to sustained anxiety, dysfunctional avoidance behaviours, and severe disability [**?**]; indeed, pain-related fear is key to the transition from acute to chronic pain.

All such disorders have a large socio-economic impact, specially after the COVID-19 pandemic outbreak [17]. Steps towards a principled understanding of FG elicited behaviours and their operationalization are thus deemed to be relevant to afford critical insight into its basic mechanisms and to shed light on viable and novel behavioural therapies [5,29].

2 Background and hypotheses

In this Section we briefly and critically overview the fear conditioning problem. This succint discussion motivates the research hypotheses behind the present work.

Fear conditioning. The behavioural mechanisms of fear learning are usually investigated using the fear-conditioning protocol in the Pavlovian tradition [26]. In the acquisition phase, one neutral stimulus (CS+), such as a light, a tone or a simple shape, is repeatedly paired with an aversive unconditioned stimulus (US), such as an uncomfortable electric shock, while another neutral stimulus is never paired with the US (CS-). After only a few CS+/US pairings, presentations of the CS+ alone will elicit a conditioned fear response (CR)

FG tests introduce generalisation stimuli (GS), which typically include several stimuli that vary in perceptual or conceptual similarity to the CS+ and are never paired with the US. generalisation of conditioned fear shows that CRs are often elicited by stimuli not associated with the aversive event but which resemble the CS+ along a perceptual or categorical dimension. FG in humans is generally measured by self-report (e.g., ratings of fear or US expectancy) or by gradients of physiological responding (e.g., skin conductance responses, SCRs, fear-potentiated startle responses, FPS, and neuro-functional activation, e.g., BOLD fMRI). Perceptual similarity is acquainted as the most accessible factor to investigate in stimulus generalisation research, as the degree of similarity can be quantified as distance between points along a continuum. Yet, beyond light tones and shapes, it has been argued [11] that real-world fear learning situations are likely to involve complex stimuli (the dog example), rather than simple unidimensional sensory cues, and that humans routinely incorporate prior conceptual knowledge to infer unobserved properties and causal structure of details surrounding an emotional event [11,3]. For instance, category-based fear conditioning has shown that production of defensive responses can be mediated by categorization processes (fear of dogs spreading to other animals or dog-related

items and contexts, but see [11] for a discussion). Knowledge about the causal structure of the world might also play an intriguing role [11]. As Rescorla put it [28], Pavlovian conditioning should be best understood as the learning of relations among events so as to allow the organism to represent its environment; in such endeavour, rephrasing Rescorla, the organism acts as a Bayesian information seeker using logical and perceptual relations among events, along with its own prior knowledge, to infer a sophisticated representation of its world. To sum up, FG based on real-world events will necessarily incorporate higher-order processes than what traditionally surmised [11].

Neurobiological bases. A large amount of work has examined brain, neural, and hormonal mechanisms underlying FG; mechanisms at the lowest level are reviewed in [2]. As to brain regions, earlier work proposed that the hippocampus matches new stimuli to a representation of the original fearful stimulus stored in memory; the degree of similarity then affects activation in downstream regions associated with fear (or safety) in amygdalar, prefrontal, and thalamic brain regions, the amygdala playing a prominent role [21]. However, the idea that the increased amygdala response combined with reduced prefrontal cortex response results in hyperarousal or enhanced affective reactivity is a too simple one [32]. A recent meta-analysis produced a more complex neural working model involving more brain regions [34]. Further, fMRI results by Onat and Buchel [25], obtained using face stimuli, indicate that FG is not passively driven by perception. Rather it is an active process that, besides perceptual similarity, integrates threat identification and ambiguity-based uncertainty to orchestrate a flexible, adaptive fear response. Dysregulation is likely to arise due to deficiency in "top-down" control of emotional response (with a key involvement of the anterior insular cortex) while integrating atypical "bottom-up" detection or appraisal of emotional triggers [32,15]. All together, these results speak for an alternative theoretical framework that is best accomodated within the psychological construction approach where basic ingredients of the mind, namely the large-scale distributed networks within the human brain, interact so that fear dysregulation emerges in all its variety and complexity [32,15,24,4].

Research hypotheses. The above considerations suggest that, overall, FG is the result of a complex, active process peculiar to the individual and, on the methodological side, the experimental shift to ecological social stimuli such as faces is not an innocent one. Under such circumstances, it is to be expected that, in the endeavour of orchestrating a flexible and adaptive fear response, the vibrant entanglement of threat identification, ambiguity-based uncertainty, perceptual similarity and prior conceptual knowledge, moulds the individual into an active information-seeker. To some extent, Reutter and Gamer experiments [29] and Onat and Buchel analyses [25] substantiate Rescorla's original assumption at the behavioural and neural level, respectively. Onat and Buchel themselves suggest that generalisation is a fundamental ability of humans facing the hurdles in the maze of social context and takes root in foraging animals striving to find food more efficiently [25]. FG, beneath its maladaptive limit, is in the service of the

organism to gain insight into the world's events relying upon its own history and skills so to prepare defensive behaviours when encountering situations that were previously experienced as dangerous.

In this perspective, the deployment of visual attention through gaze offers an effectual window on the individual's information-seeking, foraging behaviour (see [6] for an in-depth discussion). Under the visual foraging hypothesis, we assume that a foraging-based analysis of gaze deployment over ecological stimuli is an appropriate approach to characterise the individual's visual information-seeking behaviour within a fear conditioning setting.

More precisely, in such setting we expect that, by inferring model's parameters from participants' eye-tracking data, a compact set of descriptors is provided that can be exploited to weigh latent factors behind the individual's FG process, in particular his/her social anxiety traits.

3 Method

We operationalise and assess our assumptions by considering the Reutter and Gamer FG experiment [29] and their publicly available data, in particular eye-tracking recordings and questionnaires gathered from participants. In what follows for completeness sake, we summarise those aspects of their work relevant for our analyses (Sec. 3.1), leaving to [29] and the accompanying online repository[3] for details.

As to the analyses, a number of works have recently considered eye movements modelling from the foraging perspective (e.g., [23,8,7,22,20,9]) or the closely related one that exploits sequential decision-making based on drift-diffusion models (e.g. [10,31]). With respect to the analyses needed here such models are overly complex, given their aim of actually simulating gaze shifts, and/or just suitable to cope with simple visual stimuli. Thus, we draw to some extent on the approach proposed in [14] (recapped in Section 3.2), which provides a succinct phenomenological model of gaze behaviour adequate to our analyses.

Furthermore, we utilize this gaze model to extract a characterization of the participants, which is subsequently employed for the classification of their Social Interaction Anxiety Scale (SIAS) scores, providing insights into the relationship between gaze behaviour and social anxiety levels.

3.1 Participants and procedure

The main hypothesis was that participants who deploy more attention toward diagnostic facial features would also show fear generalisation gradients that are curved more strongly, thus indicating less generalisation.

Participants The participants' sample consisted of 44 individuals (35 female; age $= 25.7 \pm 5.0$ years) that exhibited low levels of social anxiety (SIAS $= 19.7 \pm 8.0$).

[3] https://osf.io/wgqnj/

Stimuli Edited facial portrait photographs (642 × 676 pixels resolution) with neutral expression were used as stimuli. Pairs of photographs with the same gender and similar skin color were selected and edited such that either the eyes or both mouth and nose were copied from one face to the other while keeping all other pixels identical. The original and the edited picture differed only in the specified area and were later used as CS+ and CS- This factor, called the diagnostic region, calls for a flexible strategy of facial exploration in order to successfully distinguish between different pairs of stimuli. In total, one male and one female pair was created for each diagnostic region (eyes vs. mouth/nose). Morphs in steps of 20% between CS+ and CS- were generated to test FG.

Procedure Subjects were seated inside an acoustically shielded cabin and performed a discrimination task to assure that they were able to perceptually distinguish between the stimuli that were sought to be used as CSs. Electrodes for pain application were attached. The intensity of the painful stimulus was then individually adjusted to a moderate level (6 or 7 on an 11-point scale ranging from 0 = no sensation over 4 = minimally painful to a theoretical maximum of 10 = worst pain imaginable). A training task ensured that subjects performed their trial-by-trial shock expectancy rating within the specified time window.

The FG task consisted of three phases. 1) *Habituation*: each of the four selected facial stimuli (one male and one female pair differing in the eye or mouth/nose region, respectively, in counterbalanced fashion) was presented four times without any electrotactile stimulation. 2) *Acquisition*: followed up seamlessly consisting of 32 trials in total. The two stimuli that were assigned to denote the CS+s (i.e., one stimulus of the male and female pair, respectively, in counterbalanced fashion) were reinforced in 75% of the cases. After a short break and recalibration of the eye-tracker, the last phase started. 3) *generalisation*: four intermediate stimuli between the CS+ and the CS- (morphs in steps of 20%) were also presented for each stimulus pair. These GS spanned the generalisation continuum from CS- across GS1 through GS4 to CS+. Each of the ten non-CS+s was presented eight times during the generalisation phase. The two CS+s, however, were presented 16 times with a 50% reinforcement rate in order to reduce extinction. Thus, the generalisation phase consisted of 112 trials. The FG task in total included 160 trials.

Each trial consisted of a face being presented for 6 seconds. After 4 secs, a rating prompt appeared at the bottom of the screen for 2 secs. Subjects were asked to indicate the perceived likelihood of an electrotactile stimulation occurring at the end of the trial on a 5-point scale (1 = no shock, 3 = uncertain, 5 = shock certain) but they were not told about different phases or contingencies. The painful stimulation was applied or omitted 5.85 secs after stimulus onset.

Eventually, subjects completed the SIAS and a demographic data questionnaire.

Data recording Trial-by-trial shock expectancy ratings (via keyboard), heart rate, pupillary responses, electrodermal activity, and eye movements were mea-

sured. Eye movements and pupil size were recorded with an EyeLink 1000 Plus system (1000 Hz)

3.2 Proposed model and data analyses

Use the the time-varying location vector $\mathbf{x}(t)$ (screen coordinates) to denote the gaze position at time t. Each observed trajectory $\{\mathbf{x}(t),\ t = 0 \dots T\}$ is a realization of a stochastic process $\{X(t),\ t = 0 \dots T\}$, with $X(t) = \mathbf{x}(t)$. A compact description of both local fixational gaze movements, occurring within a selected region of the visual field, and saccadic relocations between regions can be given in terms of a particle randomly wandering but being pulled towards an attractor (the center of a region of interest, RoI) can be formalised as a switching Ornstein-Uhlenbeck (O-U) process. To such end, denote $s_t \in [fix, sac]$, a switching state random variable, indicating whether at time t either a fixational movement within a RoI or a saccadic relocation between RoIs is performed by the observer [10][8]. Then, the stochastic differential equation (SDE)

$$d\mathbf{x}(t) = \mathbf{B}^{s_t}(\boldsymbol{\mu}^{s_t} - \mathbf{x}(t))dt + \boldsymbol{\Gamma}^{s_t} d\mathbf{W}^{s_t}(t) \tag{1}$$

accounts for the switching O-U process dynamics within and between RoIs sequentially selected by the observer as foci of attention. The term $\mathbf{B}^{s_t}(\boldsymbol{\mu}^{s_t} - \mathbf{x}(t))$ represents the drift towards the attractor point $\boldsymbol{\mu}^{s_t}$, where the 2×2 matrix $\mathbf{B}^{s_t}$ $\mathbf{B}^{s_t} = \begin{bmatrix} B_{ii}^{s_t} & B_{ij}^{s_t} \\ B_{ji}^{s_t} & B_{jj}^{s_t} \end{bmatrix}$ controls the magnitude of the attraction effect; B_{ii} and B_{jj} represent the drift of the process towards the attractor in the i (horizontal) and j (vertical) dimensions, respectively, while the off-diagonal elements $B_{ij} = B_{ji} = \rho_B \sqrt{B_{ii} B_{jj}}$ encode the cross-correlation between drift in both dimensions. The stochastic term $\boldsymbol{\Gamma}^{s_t} d\mathbf{W}^{s_t}(t)$ accounts for diffusion. Akin to $\mathbf{B}^{s_t}$, the 2×2 matrix $\boldsymbol{\Gamma}$ is the control parameter (variances and covariances) of the two driving white noise processes (horizontal and vertical) described by $d\mathbf{W}(t)$. Higher values of variances/covariances generate noisier/more anisotropic gaze trajectories. Given the set of parameters $\boldsymbol{\theta} = \{s_t, \mathbf{B}^{s_t}, \boldsymbol{\Gamma}^{s_t}, \boldsymbol{\mu}^{s_t}\}$, the simulation of a sequence of eye movements $\mathbf{x}(t) \rightarrow \mathbf{x}(t')$, with $t' > t + \delta t$, δt being an arbitrary time step, can be obtained by solving Equation 1. In generative form, the solution can be written as the conditional sampling of $\mathbf{x}(t')$ given $\mathbf{x}(t)$, i.e., $\mathbf{x}(t') \mid \mathbf{x}(t) \sim P(\mathbf{x}(t') \mid \mathbf{x}(t))$, where the distribution $P(\cdot)$ is the Normal distribution $\mathcal{N}(\cdot)$ (see e.g. [18]):

$$\mathbf{x}(t') \mid \mathbf{x}(t) \sim \mathcal{N}(\boldsymbol{\mu}^{s_t} + e^{-\mathbf{B}^{s_t}\delta t}(\mathbf{x}(t) - \boldsymbol{\mu}^{s_t}), \boldsymbol{\Psi}^{s_t}), \tag{2}$$

where $\boldsymbol{\Psi} = \boldsymbol{D}^{s_t} - e^{-\mathbf{B}^{s_t}\delta t}\boldsymbol{D}^{s_t} e^{-\mathbf{B}^{s_t T}\delta t}$; $\mathbf{B}^{s_t}$ and $\boldsymbol{D} = \frac{\boldsymbol{\Gamma}^2}{2}\mathbf{B}^{-1}$ are 2×2 matrices and the form $e^{-\mathbf{M}}$ denotes the matrix exponential.

The set of model parameters $\boldsymbol{\theta} = \{s_t, \mathbf{B}^{s_t}, \boldsymbol{\Gamma}^{s_t}, \boldsymbol{\mu}^{s_t}\}$ gives a complete description of gaze dynamics and can be inferred as follows. First, the raw eye-tracking data of an individual's gaze trajectory is parsed via the NSLR-HMM algorithm [27], to provide a set of fixational events (saccades are here discarded).

Only the fixations dwelling inside the diagnostic regions, i.e. eyes vs. mouth/nose (cfr Section 3.1) are retained. Call $\mathbf{e}^{eye} = [e_1, ..., e_{F_1}]$ the ensemble of F_1 fixations inside the eyes diagnostic area and $\mathbf{e}^{mn} = [e_1, ..., e_{F_2}]$ the group of F_2 fixations inside the mouth/nose diagnostic area. Define $\boldsymbol{\xi} = [\mathbf{e}^{eye}|\mathbf{e}^{mn}]$.

Consider the slice $\mathbf{x}^e = [\mathbf{x}_m, ..., \mathbf{x}_q]$ of the sample $\mathbf{x}(t)$, with $m \geqslant 0$ and $q \leqslant n$; the e index represents a generic fixation $e \in \boldsymbol{\xi}$. The likelihood of the slice, given the parameters $\{\mathbf{B}^e, \boldsymbol{\Gamma}^e\}$ writes $P(\mathbf{x}^e \mid \mathbf{B}^e, \boldsymbol{\Gamma}^e) = \prod_{i=1}^{q-m-1} P(\mathbf{x}_{i+1}^e \mid \mathbf{x}_i^e, \mathbf{B}^e, \boldsymbol{\Gamma}^e)$. Then, the posterior probability of the O-U parameters of the event e given the gaze trajectory slice is recovered via Bayes' theorem

$$P(\mathbf{B}^e, \boldsymbol{\Gamma}^e \mid \mathbf{x}^e) = \frac{P(\mathbf{x}^e \mid \mathbf{B}^e, \boldsymbol{\Gamma}^e) P(\mathbf{B}^e, \boldsymbol{\Gamma}^e)}{P(\mathbf{x}^e)}, \tag{3}$$

where under the mean field approximation $P(\mathbf{B}^e, \boldsymbol{\Gamma}^e) \approx P(\mathbf{B}^e)P(\boldsymbol{\Gamma}^e)$, the LKJ distribution is adopted as the prior for the $\mathbf{B}^e$ and $\boldsymbol{\Gamma}^e$ matrices in order to ensure all positive eigenvalues. Next, the event parameter posterior in Eq. 3 is computed in approximate form via Automatic Differentiation Variational Inference (ADVI) [19] and summarised through its sample average and uncertainty (Highest Density Interval, HDI). The distribution summaries are joined together, thus yielding the vector $\mathbf{v}_{(id)}^e$ for each subject $id \in [1, ..., ID]$, ID being the total number of subjects:

$$\mathrm{v}_{(id),k}^e = [B_{ii}^{avg,e}, B_{ii}^{hdi,e}, B_{ij}^{avg,e}, B_{ij}^{hdi,e}, B_{jj}^{avg,e}, B_{jj}^{hdi,e}, \Gamma_{ii}^{avg,e}, \Gamma_{ii}^{hdi,e}, \Gamma_{ij}^{avg,e}, \Gamma_{ij}^{hdi,e}, \Gamma_{jj}^{avg,e}, \Gamma_{jj}^{hdi,e}]. \tag{4}$$

Eventually, the sequence of events (fixations) - each event e being summarised by the vector $\mathbf{v}_{(id),k}^e$-, characterises the visual behaviour of observer id while scrutinising the stimulus k (image).

Denote:

- $\langle \mathbf{v}_{(id),k}^{eye} \rangle$ and $\langle \mathbf{v}_{(id),k}^{mn} \rangle$ the average fixation feature vector relative to either the eye or mouth/nose diagnostic region associated to the scan path (image) k:

$$\langle \mathbf{v}_{(id),k}^{e^{eye}} \rangle = \frac{1}{F_1} \sum_{a=1}^{F_1} \mathbf{v}_{(id),k}^{e_a^{eye}}, \qquad \langle \mathbf{v}_{(id),k}^{e^{mn}} \rangle = \frac{1}{F_2} \sum_{a=1}^{F_2} \mathbf{v}_{(id),k}^{e_a^{mn}} \tag{5}$$

- $\mathbf{v}_{(id),k}$ the descriptor of scan path k obtained by concatenating the two vectors above:

$$\mathbf{v}_{(id),k} = \left[\langle \mathbf{v}_{(id),k}^{eye} \rangle | \langle \mathbf{v}_{(id),k}^{mn} \rangle\right]; \tag{6}$$

- $\langle \mathbf{v}_{(id)} \rangle$ the summary descriptor of the visual behaviour of observer id, over the set of the K observed stimuli:

$$\langle \mathbf{v}_{(id)} \rangle = \frac{1}{K} \sum_{k=1}^{K} \mathbf{v}_{(id),k}. \tag{7}$$

The categorization of each fixation to its eventual diagnostic region has been carried out utilizing the pre-existing masks within the dataset, which were employed for the purpose of the original study. The account for that was motivated

by the potential diagnostic value of these facial regions. By dividing fixations into these two categories, we sought to capture the nuanced dynamics of visual attention within different facial regions and explore their specific contributions to the prediction of social anxiety levels. The extracted parameters and features were derived separately for fixations related to the eyes and fixations related to the region of the nose/mouth, enabling a more fine-grained analysis of the observer's gaze behaviour and its potential relevance for social anxiety prediction. The extracted gaze dynamics parameters were then used as inputs for a Random Forest (RF) classifier to predict the SIAS score. To simplify the prediction task and facilitate the interpretation of results, we transformed the social anxiety prediction problem into a binary classification task. Given that the SIAS scores range from 0 to 80, we decided to use the median score in our dataset, which was found to be 18, as the threshold for binarizing the scores. Individuals with SIAS scores equal to or above the median threshold were considered to have high social anxiety, while those below the threshold were classified as having low social anxiety. The transformation into a binary classification problem allowed us to utilize well-established classification algorithms, such as the Random Forest classifier, to predict social anxiety levels effectively. Prior to selecting the RF classifier, we conducted an evaluation of other classification algorithms, such as the Support Vector Machine (SVM) with a radial basis function kernel and the linear Support Vector Machine classifier (linSVM).

4 Results

We utilized data from a cohort of 43 participants out of the initial pool of 44 participants, as recordings from one participant resulted to be inoperable and had to be excluded from our study.

The evaluation process involved 5-fold cross-validation to ensure robustness and mitigate any potential biases. Performance was assessed using accuracy.

The results of the evaluation revealed that the RF classifier outperformed the other algorithms with a significantly higher accuracy score of 0.73. In comparison, the linSVM and SVM algorithms yielded lower accuracy scores of 0.58 and 0.61 respectively 1.

Algorithm	Accuracy (5-fold cv)
linSVM	0.58
SVM	0.61
RF	**0.73**

Table 1. Accuracy results of tested algorithms for SIAS classification.

Additionally, we conducted an analysis to assess the feature importance, computed as the mean of accumulation of the impurity decrease within each

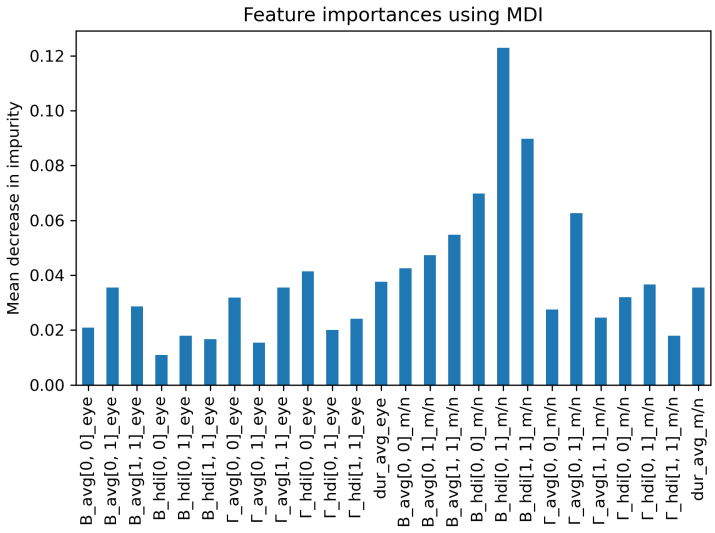

Fig. 2. Results of feature importance estimate, where indexes in brackets indicate the specific element of the matrix.

tree (Fig. 2). This analysis provided valuable insights into the significance of different gaze dynamics features for the task at hand.

Notably, the results of this analysis revealed that features related to the gaze dynamics of the nose/mouth region hold greater importance in predicting social anxiety levels. Specifically, the feature that exhibited particularly high importance was the standard deviation of the magnitude of fixations drift towards the mean (B matrix).

5 Conclusions

In this paper, we have presented a systematic and principled approach for analyzing eye movements in the context of fear generalisation, specifically focusing on the prediction of Social Interaction Anxiety Scale (SIAS) scores.

In the framework of foraging theory applied to eye movements, we have introduced a composite O-U process to operationalise social anxiety assessment. This phenomenological model captures the exploration-exploitation signature inherent in foraging eye behaviour. By inferring the relevant parameters of the composite O-U model through Bayesian analysis of eye-tracking data, we have identified a feature set that is suitable for predicting SIAS scores.

The results of our study demonstrate the effectiveness of the proposed approach. By utilizing the inferred parameters from the composite O-U model of fixations as features, we have achieved promising performance in predicting SIAS scores using Random Forest (RF) as classification technique.

This research contributes to the body of knowledge by providing a novel perspective on the analysis of gaze for social anxiety assessment. By embracing a model-based approach and leveraging principles from foraging theory, this work can open novel avenues for future research in understanding and utilizing gaze behaviour as a valuable modality for psychological and behavioural assessment.

References

1. Ahrens, L.M., Pauli, P., Reif, A., Mühlberger, A., Langs, G., Aalderink, T., Wieser, M.J.: Fear conditioning and stimulus generalization in patients with social anxiety disorder. Journal of Anxiety Disorders **44**, 36–46 (2016)
2. Asok, A., Kandel, E.R., Rayman, J.B.: The neurobiology of fear generalization. Frontiers in behavioral neuroscience **12**, 329 (2019)
3. Barrett, L.F.: The theory of constructed emotion: an active inference account of interoception and categorization. Social cognitive and affective neuroscience **12**(1), 1–23 (2017)
4. Barrett, L.F.: Seeing fear: it's all in the eyes? Trends in neurosciences **41**(9), 559–563 (2018)
5. Beckers, T., Hermans, D., Lange, I., Luyten, L., Scheveneels, S., Vervliet, B.: Understanding clinical fear and anxiety through the lens of human fear conditioning. Nature Reviews Psychology pp. 1–13 (2023)
6. Bella-Fernández, M., Suero Suñé, M., Gil-Gómez de Liaño, B.: Foraging behavior in visual search: A review of theoretical and mathematical models in humans and animals. Psychological research **86**(2), 331–349 (2022)
7. Błażejczyk, P., Magdziarz, M.: Stochastic modeling of lévy-like human eye movements. Chaos (Woodbury, NY) **31**(4), 043129 (2021)
8. Boccignone, G., Cuculo, V., D'Amelio, A., Grossi, G., Lanzarotti, R.: On gaze deployment to audio-visual cues of social interactions. IEEE Access **8**, 161630–161654 (2020)
9. Boccignone, G., Cuculo, V., D'Amelio, A., Grossi, G., Lanzarotti, R.: Give ear to my face: modelling multimodal attention to social interactions. In: Proceedings of the European Conference on Computer Vision (ECCV) Workshops. pp. 0–0 (2018)
10. D'Amelio, A., Boccignone, G.: Gazing at social interactions between foraging and decision theory. Frontiers in neurorobotics **15**, 31 (2021)
11. Dunsmoor, J.E., Murphy, G.L.: Categories, concepts, and conditioning: how humans generalize fear. Trends in cognitive sciences **19**(2), 73–77 (2015)
12. Dunsmoor, J.E., Paz, R.: Fear generalization and anxiety: behavioral and neural mechanisms. Biological psychiatry **78**(5), 336–343 (2015)
13. Dymond, S., Dunsmoor, J.E., Vervliet, B., Roche, B., Hermans, D.: Fear generalization in humans: systematic review and implications for anxiety disorder research. Behavior therapy **46**(5), 561–582 (2015)
14. D'Amelio, A., Patania, S., Bursic, S., Cuculo, V., Boccignone, G.: Using gaze for behavioural biometrics. Sensors **23**(3), 1262 (2023)
15. Fitzgerald, J.M., DiGangi, J.A., Phan, K.L.: Functional neuroanatomy of emotion and its regulation in ptsd. Harvard review of psychiatry **26**(3), 116 (2018)
16. Ghirlanda, S., Enquist, M.: A century of generalization. Animal Behaviour **66**(1), 15–36 (2003)
17. Kar, N., Kar, B., Kar, S.: Stress and coping during covid-19 pandemic: Result of an online survey. Psychiatry research **295**, 113598 (2021)

18. Kloeden, P.E., Platen, E.: Numerical solution of stochastic differential equations, vol. 23. Springer Science & Business Media, Berlin, Germany (2013)
19. Kucukelbir, A., Tran, D., Ranganath, R., Gelman, A., Blei, D.M.: Automatic differentiation variational inference. The Journal of Machine Learning Research **18**(1), 430–474 (2017)
20. Le, S.T.T., Kristjansson, A., MacInnes, W.J.: Bayesian approximations to the theory of visual attention (tva) in a foraging task. Quarterly Journal of Experimental Psychology **76**(3), 497–510 (2023)
21. Lissek, S., Bradford, D.E., Alvarez, R.P., Burton, P., Espensen-Sturges, T., Reynolds, R.C., Grillon, C.: Neural substrates of classically conditioned fear-generalization in humans: a parametric fmri study. Social cognitive and affective neuroscience **9**(8), 1134–1142 (2014)
22. Mirza, M.B., Adams, R.A., Friston, K., Parr, T.: Introducing a bayesian model of selective attention based on active inference. Scientific reports **9**(1), 1–22 (2019)
23. Mirza, M.B., Adams, R.A., Mathys, C.D., Friston, K.J.: Scene construction, visual foraging, and active inference. Frontiers in Computational Neuroscience **10**, 56 (2016), https://www.frontiersin.org/article/10.3389/fncom.2016.00056
24. Mobbs, D., Adolphs, R., Fanselow, M.S., Barrett, L.F., LeDoux, J.E., Ressler, K., Tye, K.M.: Viewpoints: Approaches to defining and investigating fear. Nature neuroscience **22**(8), 1205–1216 (2019)
25. Onat, S., Büchel, C.: The neuronal basis of fear generalization in humans. Nature neuroscience **18**(12), 1811–1818 (2015)
26. Pavlov, I.P.: Conditioned Responses. Prentice-Hall, Inc (1949)
27. Pekkanen, J., Lappi, O.: A new and general approach to signal denoising and eye movement classification based on segmented linear regression. Scientific reports **7**(1), 1–13 (2017)
28. Rescorla, R.A.: Pavlovian conditioning: It's not what you think it is. American psychologist **43**(3), 151 (1988)
29. Reutter, M., Gamer, M.: Individual patterns of visual exploration predict the extent of fear generalization in humans. Emotion (September 2022), https://doi.org/10.1037/emo0001134
30. Roesmann, K., Wiens, N., Winker, C., Rehbein, M.A., Wessing, I., Junghoefer, M.: Fear generalization of implicit conditioned facial features–behavioral and magnetoencephalographic correlates. Neuroimage **205**, 116302 (2020)
31. Roth, N., Rolfs, M., Hellwich, O., Obermayer, K.: Objects guide human gaze behavior in dynamic real-world scenes. bioRxiv (2023), https://www.biorxiv.org/content/early/2023/03/14/2023.03.14.532608
32. Suvak, M.K., Barrett, L.F.: Considering ptsd from the perspective of brain processes: A psychological construction approach. Journal of traumatic stress **24**(1), 3–24 (2011)
33. Tenenbaum, J.B., Kemp, C., Griffiths, T.L., Goodman, N.D.: How to grow a mind: Statistics, structure, and abstraction. Science **331**(6022), 1279–1285 (2011)
34. Webler, R.D., Berg, H., Fhong, K., Tuominen, L., Holt, D.J., Morey, R.A., Lange, I., Burton, P.C., Fullana, M.A., Radua, J., et al.: The neurobiology of human fear generalization: meta-analysis and working neural model. Neuroscience & Biobehavioral Reviews **128**, 421–436 (2021)