ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

SCIENZE CHIRURGICHE E TECNOLOGIE INNOVATIVE

Ciclo 38

Settore Concorsuale: 11/E1 - PSICOLOGIA GENERALE, PSICOBIOLOGIA E PSICOMETRIA

Settore Scientifico Disciplinare: M-PSI/01 - PSICOLOGIA GENERALE

VISUALIZING THE MENTAL REPRESENTATION OF A "GOOD DOCTOR": TESTING ITS POTENTIAL EFFECTS ON PAIN MODULATION

Presentata da: Ilenia Ceccarelli

Coordinatore Dottorato

Emanuela Marcelli

Supervisore

Katia Mattarozzi

Co-supervisore

Arianna Bagnis

Esame finale anno 2026

Abstract

This PhD project explores the influence of social stimuli, particularly facial cues, on pain perception and the potential for placebo hypoalgesia. The research bridges the gap between social modulation of pain and facial-appearance based inferences, aiming to understand these effects more deeply.

In the first study, a group-level mental representation of a “good doctor” was developed using ratings and machine learning with an Italian sample. This model manipulated facial features to match perceived traits— ethnicity and perceived similarity, attractiveness, trustworthiness, familiarity and professional look/intelligence —highlighting its relevance to healthcare.

The second study examined if this representation influences pain perception by reducing threat through familiarity inferences, considering genetic factors. Results showed modulation of pain intensity and unpleasantness by faces aligning with the “good doctor” representation, with genetic polymorphisms affecting the outcome.

Study 3 focused on creating idiosyncratic models of the “good doctor” stereotype using generative models. These personalized models reflected participants' own stereotypes more accurately and were seen as more familiar, emphasizing intrinsic familiarity as crucial.

The fourth study preliminarily assessed (N=30) whether idiosyncratic representations modulated pain perception more effectively than non-idiosyncratic ones. Preliminary findings, based on a small sample, confirm a significant modulation of the pain affective component, such as unpleasantness, particularly at the idiosyncratic level (between high and low representations). Overall, this project underlines the importance of facial cues as social signals influencing pain responses, contributing to theoretical insights and potentially guiding clinical interventions.

Summary

The presence of conspecifics impacts the subjective experience of pain, potentially leading to placebo hypoalgesia (Bagnis et al., 2023; Krahè et al., 2013; Mattarozzi et al., 2021). This effect may have evolved to help individuals to adaptively respond to aversive situations, such as painful experiences, in the presence of others.

As powerful social stimuli, human faces swiftly convey inferences of social traits such as familiarity, trustworthiness and dominance, shaping people's dispositions and social outcomes (Todorov et al., 2015). In the healthcare field, it has been observed that patients' perceptions of the healthcare provider's facial appearance impact their experience of medical procedures, with higher perceived familiarity associated with higher pain thresholds, lower pain intensity, and greater treatment satisfaction (Mattarozzi et al., 2021; Bagnis et al., 2023).

Across four different studies, the present PhD project bridges the gap between these two lines of investigation - the social modulation of pain and the impact of facial-appearance based inferences on behaviour- aiming to pursue four different and interrelated goals.

The first study focused on developing and validating the mental and stereotypical representation of a "good doctor" at the group level. Using ratings from an Italian sample and machine- learning methods, we generated a model capable of manipulating facial features to match this mental representation. The model correlated with models of social attributes loadings in four main dimensions: ethnicity and perceived similarity, attractiveness, trustworthiness, and familiarity and professional look/intelligence.

The second study investigated whether this stereotypical representation modulates the subjective pain experience by reducing the sense of threat through inferences of familiarity. The role of individual factors, such as the genetic profile, on this modulatory effect was also explored. Faces aligning with the mental representation of a "good doctor" modulate both pain intensity ($F(2,3218)=7.01$, $p<.001$, $R^2c=.75$, $R^2_m=.001$) and unpleasantness ($F(2,3218)=8.96$, $p<.001$, $R^2c=.72$, $R^2_m=.002$); with these effects influenced by genetic polymorphisms related to endocannabinoid, dopaminergic, serotoninergic and oxytocinergic systems.

In light of the findings on group models, Study 3 focused on creating and validating idiosyncratic (i.e., participant-specific) representations of the "good doctor" using generative models. In line with

our hypothesis, these personalized models were more accurate in representing participants' own "good doctor" stereotypes than those of others (b=-0.99, t(2046.81) =-12.04, p<.001), and were also rated as more familiar (b=-0.45, t(2077.02)=-4.07, p<.001). This study highlighted both shared and individual components of the "good doctor" stereotype, emphasizing "intrinsic familiarity" as a central trait.

Finally, the fourth study aimed to evidence whether these idiosyncratic representations of the "good doctor" stereotype may be more effective in modulating pain perception compared to non-idiosyncratic representations. Initial results, based on a limited sample, suggest a significant modulation of the affective component of pain—i.e., unpleasantness— at the idiosyncratic level (between high and low representations).

Together, the present PhD project highlights the significant role of social silent stimuli, such as facial cues, in shaping potential adaptive responses as placebo hypoalgesia. These findings contribute to a richer theoretical understanding of pain processing and may inform more targeted interventions in both pain research and clinical practice.

Index

Abstract 1

Summary 2

1. General introduction: Theoretical Background 8

1.1 The social modulation of pain 8

1.2 Social evaluation from faces: underlying mechanisms, individual differences and clinical implications 10

1.2.1 Social attributions from faces: determinants and consequences on individuals' behaviour 10

1.2.2 Idiosyncratic contributions to Face Evaluation 13

1.2.3 The role of face- based social attributions in the healthcare context 14

1.2.3.1 The effect of social attributions on pain experience 16

1.3 Individual Differences, Genetic variants and the social modulation of pain 17

2 Research projects and Objectives 22

3 A data-driven method for visualizing the "Good doctor" stereotypical representation 24

3.1 Introduction 24

3.2 Experiment 1 25

3.2.1 Materials and Methods 25

3.2.2 Results 29

3.3 Experiment 2 33

3.3.1 Materials and Methods 34

3.3.2 Results 36

3.4 Discussion 37

3.5 Conclusion 42

4 From stereotypes to pain: How faces aligning with the “good doctor” mental representation modulates pain perception 43

4.1 Introduction 43

4.2 Materials and Methods 45

4.3 Results 53

4.3.1 Pain measures: pain intensity and unpleasantness 53

4.3.2 Psychophysiological measure: Mean HR and HRV components during pain 58

4.3.3 Questionnaires 59

4.4 Discussion 63

4.5 Conclusions 68

5 Visualizing the idiosyncratic representation of a “Good doctor”: development and validation 70

5.1 Introduction 70

5.2 Experiment 1 71

5.2.1 Materials and Methods 71

5.2.2 Results 75

5.3 Experiment 2 76

5.3.1 Materials and Methods 77

5.3.2 Results 80

5.4 Discussion 81

5.5 Conclusions 85

6 The effects of idiosyncratic “good doctor” stereotypical representation on pain perception 86

6.1 Introduction 86

6.2 Materials and Methods 87

6.3 Preliminary results 96

6.3.1 Pain task 96

6.3.2 Familiarity task 97

6.3.3 Questionnaires 99

6.4 Discussion 104

6.5 Conclusions 106

7. General conclusions 108

References 112

8 Supplementary materials 144

8.1 Chapter 4: additional results 144

8.1.1 Exploratory analyses: the role of sex differences 144

8.1.2 Main effect of Condition on pain intensity and unpleasantness 150

8.1.3 Condition x Gene x Gene interaction: COMT rs4680 (AA vs AG vs GG) and FAAH rs324420 (A* vs CC) 152

8.1.4 Condition x Gene x Gene interaction: TPH2 rs4570625 (GG vs T*) and OXTR rs2254298 (CC vs T*) 156

8.1.5 Physiological measures: main effect of Condition on HRV components 160

8.1.6 Questionnaires: correlational analyses between questionnaires scorings 162

8.2 Chapter 5: Additional results 168

8.2.1 Idiosyncratic “good doctor” models: Exploratory analyses on prevalent sex and age of the idiosyncratic “good doctors” 168

8.2.2 Idiosyncratic “good doctor” models: Exploratory analyses on the social attributes underlying the idiosyncratic “good doctors” 169

8.3 Chapter 6: Additional results 171

8.3.1 Main effect of Condition on pain intensity and unpleasantness 171

8.3.2 Pain Task: Main effect of Condition on Mean HR and HRV components (i.e., RMSSD,VLF,LF,HF,LF/HF) 174

8.3.3 Familiarity task: Main effect of Condition on Familiarity ratings 179

8.3.4 Questionnaires: correlational analyses between pain measures and questionnaires scorings 181

1 General introduction: Theoretical Background

1.1 The social modulation of pain

Pain is a sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage (Raja et al., 2020). Similar to other perceptual stimuli, it arises from the integration of both bottom-up processes—reflecting its nociceptive and interoceptive components (Craig, 2002)—and top-down mechanisms, including learning, memory, expectations (Atlas & Wager, 2012), mood (Villemure & Bushnell, 2009), and cognitive appraisal (Vlaeyen et al., 2009).

From an evolutionary perspective, pain serves an adaptive function by motivating withdrawal behaviours and alerting both the individual and the nearby conspecifics to potential environmental threats (de C Williams & Craig, 2006). However, beyond this role, the social implications of pain have likely played an equally important part in fostering the evolutionary success of the human species (Sullivan & de C Williams, 2025). Expressing pain elicits caregiving behaviours and promotes empathy and cooperation among group members (de C Williams et al., 2016; de C Williams, 2023; M. D. Sullivan & de C Williams, 2025).

Adopting this perspective, pain can therefore be considered to have an intrinsic social nature (Sullivan & Williams, 2025), and its experience is deeply shaped by the social context in which it occurs (Carlino et al., 2014; Hadjistavropoulos et al., 2011; Karos et al., 2018, 2020).

Threatening social context– as those where pain is perceived to be inflicted by others intentionally or unjustly – worsens pain perception (Karos et al., 2020; Trost et al., 2014). Conversely, social support alleviates pain, as evidenced by reduced self- reported pain measures (Brown et al., 2003; Reddan et al., 2020) and related physiological arousal (Roberts et al., 2015). For instance, verbal support- especially when clearly communicated and provided by a significant other (e.g., romantic partner)- effectively reduced pain (Krahé et al., 2013). Holding the hand of a significant other led to decreased pain intensity and unpleasantness, heart rate acceleration induced by painful stimulations (Che et al., 2018; Reddan et al., 2020), and reduced experimentally induced secondary hyperalgesia (Jaltare et al., 2023).

This line of research demonstrated that providing social support without actual presence-i.e., primed support- can also modulate pain. Experimental studies suggest that mentally activating a close

relationship, for example imagining supportive touch from one's partner, can buffer pain and stress (Jakubiak & Feeney, 2016). Similarly, individuals exposed to the picture of an affective other– as partner's picture – during noxious stimulation reported decreased pain intensity (Eisenberger et al., 2011; Hillmer et al., 2021) and reduced neural activity in pain - related regions (i.e., dorsal anterior cingulate cortex, and anterior insula) (Eisenberger et al., 2011). Pictures of a romantic partner are imbued with positive value, and their presentation is sufficient to activate positive mental representation in the observer (Ferguson & Bargh, 2004). Moreover, partner's pictures are also familiar, which make them more socially relevant stimuli– and hence more salient– than non-familiar ones (Jenkins et al., 2011). Notably, the presentation of partner's picture during painful stimuli was associated with increased activity in the ventromedial prefrontal cortex (vmPFC), a region associated with safety signalling (Fullana et al., 2016; Schiller & Delgado, 2010).

Primed support has been shown to effectively modulate pain even when the individual shown in the picture was a stranger (Edwards et al., 2017). For instance, the mere exposure to both partner and happy faces of a stranger led to reduced self- reported pain intensity compared to neutral faces, angry faces and objects (Hillmer et al., 2021). Moreover, this reduction correlated with reduced corrugator muscle activity, an index of perceived aversiveness of stimuli (Lang et al., 1998).

A possible explanation for the modulatory effect that facial images exert on perceived pain may be rooted in the type of information they convey. According to the social signal value of emotion expression (Hareli & Hess, 2012), emotional facial expressions provide valuable information about both the expresser and the situation in which they are involved. Moreover, as stated by the Social Baseline Theory (SBT, Coan & Sbarra, 2015), proximity to social resources reduces the perceived distress in the face of a potential threat, as the human brain has evolved by learning that survival is enhanced by others' presence. From an evolutionary perspective, this implies that in a threatening context such as pain administration, social stimuli as faces may act as a "safety signal", signalling contextual safety and inducing a feeling of "security" that, in turn, may reduce pain (Coan et al., 2006; Krahé et al., 2013; Vlaeyen et al., 2009).

Taken together, these findings point to a placebo hypoalgesic effect, understood as a psychoneurobiological mechanism involving factors related to the individual and the context that can modulate pain perception interactive with actual nociceptive stimuli (Bagnis et al., 2025). This effect can be elicited by perceived social support (Krahé et al., 2013), whether provided directly (e.g., verbal reassurance, handholding) or merely primed (e.g., face image). Perceived social support may signal the safeness of the environment in which the impending noxious stimulation occurs. This process,

combined with the nociceptive component of the stimulation and the individual's prior beliefs and appraisal, may ultimately shape the subjective experience of pain; reflecting an evolutionary mechanism enabling individuals to adaptively respond to aversive experiences – as pain- in the presence of conspecifics. Additionally, it also underscores the significance of human faces as powerful social stimuli that can influence and modulate pain perception.

1.2 Social evaluation from faces: underlying mechanisms, individual differences and clinical implications

Human faces represent a powerful source of information about conspecifics, influencing individuals' disposition toward others (Todorov, 2017; Zebrowitz & Montepare, 2008).

Over the past two decades, research has aimed to disentangle the mechanisms underlying face-based inferential processes, investigating the correspondence between physical traits, social traits and social stereotypes across various domains, including healthcare. These studies highlight not only the perceptual basis of social evaluations from faces, but also, and more importantly, their consequential impact on behavior.

Across four subsections, the present section explores this phenomenon and its involvement in pain modulation in greater depth. First, it examines how social attributions from faces emerge, with a focus on the relationship between facial features, trait inferences, and behavioral consequences. Second, it considers the role of idiosyncratic judgments in shaping these inferences. Third, it reviews the implications of face-based evaluations within clinical and healthcare settings. Finally, it links these evaluations to the broader framework of the social modulation of pain, investigating how social attributions from face may influence individuals' pain experiences.

1.2.1 Social attributions from faces: determinants and consequences on individuals' behaviour

People spontaneously attribute a wide range of characteristics from a person's face. A hundred millisecond exposure to an unfamiliar face is sufficient to generate specific impressions, and additional time seems just to strengthen—rather than alter—these initial judgments (Todorov et al., 2015). These impressions are shaped not only by variant facial features, such as emotional expressions—happy faces are typically judged as more trustworthy than angry ones (Oosterhof & Todorov, 2009)- but also by invariant features - nose size and chin pointedness relates with judgments of intelligence (Kleisner et al., 2014).

Notably, global facial configurations contribute to social attributions as well. Adults with baby-faced features are perceived as trustworthy, kind, and warm (Zebrowitz et al., 2012; Zebrowitz & Montepare, 1992). Attractive individuals are considered intelligent (Dion et al., 1972; Dion, 2002; Kanazawa, 2011), sociable and vain (Rougier et al., 2023). Finally, self-resembling faces are typically judged as more trustworthy and elicit more prosocial behaviours compared to non-self-resembling faces (DeBruine, 2002; Krupp et al., 2008); probably due to a reattribution process which makes individuals more likely to recognize themselves in trustworthy than untrustworthy morphs (Verosky & Todorov, 2010).

Although unreliable and based solely on physical appearance, social attributions from faces have consequential effects- even in contexts where more information is available (Olivola et al., 2018) or when reliance on physical appearance is discouraged (Jaeger et al., 2020). Their influence spans a variety of domains, from experimental settings (Tingley, 2014) to real-world outcomes such as electoral success (Todorov et al., 2005), hiring (Bagnis et al., 2021) and economic decisions (Van 't Wout & Sanfey, 2008). In laboratory studies, for example, participants involved in strategic economic games are less willing to trust individuals with untrustworthy-looking faces—even when those individuals' past behaviour contradicts such judgments (Rezlescu et al., 2012). Similarly, in legal contexts, defendants perceived as untrustworthy or stereotypically "fitting" the crime they are accused of are more likely to receive guilty verdicts—even when the evidentiary basis is weak (Dumas & Testé, 2006; Porter et al., 2010). Competent-looking individuals, regardless of actual performance, are more likely to be selected for leadership positions (Graham et al., 2017; Ling et al., 2019; Stoker et al., 2016).

Such face-based attributions can impact judgments in healthcare settings, influencing diagnostic and treatment decisions—a topic that will be further explored in Section 1.2.3. In addition to influencing decision-making, social attributions from facial appearance also affect emotional processes, such as recognition of emotional expressions. Although distinct from judgments that guide decisions, emotion recognition can be biased by facial appearance cues, which may facilitate or impair accurate decoding of emotional expression, with implications for social interaction and clinical assessments (Bagnis et al., 2024; Colonnello et al., 2019).

A compelling explanation for why these facial inferences have such an impact on individual's behavior may be rooted in the overgeneralization hypothesis (Zebrowitz et al., 2003). This account suggests that individuals tend to attribute specific social traits – stereotypically related to social categories, identities or emotional expressions– to people whose faces merely resemble those

categories, identities or expressions. Based on an adaptive perspective, certain facial cues signal social attributes which are so evolutionarily relevant that even subtle traces can trigger a response (Zebrowitz & Montepare, 2006). Consequently, this leads individuals to react in ways that are presumed to be adaptive, based on the inferred traits—even when those inferences are not supported by reliable information. This evolutionary interpretation is corroborated by findings evidencing that faces are evaluated on two fundamental dimensions, namely trustworthiness, which is the closest approximation of the valence of a face; and dominance (Oosterhof & Todorov, 2008),which is more strictly related to achievement drive and aggression (Cheng et al., 2013). Whereas trustworthiness evaluations signal whether to approach or avoid a person, dominance evaluations inform on their physical strength, giving rise to inferences about the person's intentions as well their ability to implement them. It derives that the errors stemming from these overgeneralizations may be less costly, adaptively, than failing to respond to potential threats or social opportunities.

Literature indicates that several moderating factors influence how these attributions shape behaviour. Among these, social context seems to play a role in determining which social traits are more influential in shaping behavior or decision-making. Within politics, for instance, perceived competence strongly predicts electoral success (Castelli et al., 2009; Olivola & Todorov, 2010); while in the military the professional success seems to be more influenced by perceived dominance (Re et al., 2013). The target's characteristics also matter: baby-facedness positively predicts success for Black male CEOs but negatively for White male CEOs (Livingston & Pearce, 2009). Furthermore, the availability of information can reduce– though not eliminate– the effect of facial impressions (Chang et al., 2010; Rezlescu et al., 2012). Finally, also the perceiver's own characteristics shape the impact these impressions have on behaviour. Faces associated with members of a social category are evaluated according to the personal beliefs about that category (Bagnis et al., 2019). Moreover, in contexts where observers have specific expertise, facial evaluations are less likely to be driven by initial impressions and more influenced by informed, experience-based considerations (Todorov et al., 2015).

To summarize, social attributions evolved from adaptive mechanisms to enhance survival (Oosterhof & Todorov, 2008; Zebrowitz et al., 2007; Zebrowitz & Montepare, 2006), and depend on different sources of information– some universally shared (e.g., inferences based on masculinity/femininity), others more idiosyncratic (e.g., resemblance to significant others).

Nevertheless, most of the evidence discussed so far is based on group-level analyses, which aggregate judgments across observers to identify stimulus-driven effects that are consistent across individuals.

While this approach has been instrumental in uncovering general patterns, it has often overlooked idiosyncratic contributions to social perception. Yet, recent findings suggest that these individual-specific factors account for a substantial portion of the meaningful variance in complex trait judgments, such as trustworthiness and familiarity (Albohn et al., 2024; Martinez et al., 2020). To disentangle the mechanisms underlying face evaluation—and given their potential role in modulating the subjective experience of pain— the role of idiosyncratic contributions will be reviewed.

1.2.2 Idiosyncratic contributions to Face Evaluation

Group- level observations have long been the standard approach for investigating how individuals form social impressions of others (Oosterhof & Todorov, 2009; Peterson et al., 2022; Vernon et al., 2014). By aggregating judgments from multiple observers, it was possible to observe how specific facial features could influence judgments on specific social attributes. Moreover, through the application of data-driven computational approaches to face evaluation, it was possible to manipulate facial images along those attributes with more specificity; enabling researchers to examine how systematic variations in facial features predict judgments along the targeted social dimension (Peterson et al., 2022).

While group- level observations have been instrumental in mapping facial features to specific social judgments (Oosterhof & Todorov, 2008; Peterson et al., 2022), they overlook a crucial source of variability: individual differences. Accumulating evidence demonstrates that individual factors such as personality, emotional disposition, age, gender, and political ideology influence how people interpret and weight facial features when inferring social attributions (Albohn et al., 2022, 2022, 2024; Arce et al., 2009; Hehman et al., 2017; Malloy et al., 2023; Mattarozzi et al., 2015). Therefore, relying solely on group-level data can be misleading— especially when the object of investigation involves complex judgments such as those derived from observing another person's face (Albohn et al., 2025; Martinez et al., 2020).

Studies partitioning social judgment variance into shared (i.e., explained by the physical features of the stimulus) and idiosyncratic components (i.e. due to the individual differences and how these differences interact with the physical features of the stimulus) have indeed revealed that idiosyncratic variance often exceed shared variance (Albohn et al., 2022, 2024; Martinez et al., 2020). Nevertheless, the degree to which idiosyncratic contributions outweigh shared ones depends on the specific social judgment being considered.

Judgments of masculinity and femininity, for instance, are largely influenced by stimulus- level, shared contributions, as they are clearly mapped onto physical facial features on which observers tend to agree (e.g. prominent eyebrows and a square jaw are typically considered masculine features (Mogilski & Welling, 2018)). In contrast, judgments of trustworthiness and familiarity are primarily explained by individual- level, idiosyncratic contributions, as there is less agreement on which facial features convey such information and how these are weighted in formulating judgments (Albohn et al., 2024, 2025). In this context, familiarity refers to the perceived resemblance to someone the observer has previously encountered or knows personally, rather than episodic familiarity. Similar findings have been observed even for social category judgments, such as sexual orientation and political ideology (Bjornsdottir et al., 2022).

These studies raise important methodological concerns regarding the use of group-level observations to investigate how individuals perceive others. What is typically discarded as noise may, in fact, mask meaningful individual-level variance that can offer valuable insights into the mechanisms by which people form social judgments from faces. In this regard, recent studies attempted to develop models of social attribute dimensions based on individuals' judgments, in order to evidence their idiosyncrasies (Albohn et al., 2022, 2025). Using data-driven methods that leverage generative artificial intelligence, these models allow the generation of psychologically aligned visual representations of social judgments at the individual level. These models have been shown to accurately predict individuals' judgments (Albohn et al., 2025) and thus represent a powerful tool to investigate how face-based social attributions influence behaviour—particularly in contexts where first impressions may have critical, real-world consequences. One of the most relevant of these is the healthcare setting, where first encounters between patients and healthcare providers are especially consequential.

1.2.3 The role of face- based social attributions in the healthcare context

As illustrated previously (for more details, see subchapter 1.2.1), social attributions from faces have been shown to predict significant outcomes across diverse domains—from politics (Todorov et al., 2005) to economics (Van 't Wout & Sanfey, 2008), and even in fields where reliance on physical appearance is explicitly discouraged, such as the juridical context (Jaeger et al., 2020). In the healthcare field, the implications of social attributions are equally notable, particularly concerning healthcare provider- patient interactions. For example, Mattarozzi et al., 2017 found that healthcare providers are biased by patients' facial appearance, as higher caring inclination scores were associated with patients' faces perceived as trustworthy compared to untrustworthy-looking faces. Interestingly,

expert healthcare providers showed higher scores than novices when assessing untrustworthy-looking faces, indicating that experience may modulate these biases. Bagnis et al., 2020 demonstrated that faces of patients perceived as more trustworthy and distressed were associated with higher priority codes in emergency settings, highlighting how facial appearance-based inferences influence the attribution of priority in situations requiring rapid decisions and limited clinical information. Furthermore, perceived trustworthiness appears to affect doctors' judgment about the authenticity and intentions of patients (Schäfer et al., 2016). Specifically, women with low trustworthy-looking faces were judged as experiencing less pain and as being more likely to exaggerate their symptoms. Lastly, trustworthiness emerged as influencing medical students' proneness to recognize pain (Ceccarelli et al., 2024). Ceccarelli et al. (2024) found that pain recognition was faster for untrustworthy-looking White faces compared to trustworthy faces; however, perceived trustworthiness had a minimal effect on recognition speed for Black faces, though untrustworthy-looking faces were estimated to express more pain, especially for Black individuals. These findings underscore how facial social attributions can significantly impact clinical judgments and behaviors in healthcare.

Patients' social attributions from healthcare providers' appearance proved to be similarly consequential. However, research on the impact of physicians' appearance on patients' impressions is still in its early stages; and has mainly focused on isolated aspects of doctors' appearance—such as posture (Grün et al., 2022) and dress style (Rehman et al., 2005)—or on broader psychosocial features, such as ethnicity and sex (Hall et al., 2020). For instance, it has been observed that physicians assuming open rather than closed postures were generally rated as more competent (Grün et al., 2022). Moreover, doctors wearing the white coat were rated higher in competence, trustworthiness and empathy compared to doctors wearing other attire (Chung et al., 2012). Sex and age also shape patient's first impressions of doctors: female doctors are perceived as less competent than male colleagues (Hall et al., 2020), with perceived competence increasing with age—a trend especially evident among women. Conversely, female physicians tend to be evaluated as more patient- centered than males (Hall et al., 2020; Roter & Hall, 2004), whereas patient-centeredness declines with age in men- suggesting how younger male doctors tend to be viewed as departing from the traditional, authoritative model of older physicians (Hall et al., 2020). Furthermore, perceived similarity to the self seem to correlate with greater intent to adhere to prescribed treatment (Necka et al., 2021; Street et al., 2008). Finally, smiling exerts a particularly robust effect as well, with smiling doctors being perceived as more competent and, most strikingly, more desirable to consult again (Hall et al., 2020).

Taken together, these findings underscore the crucial role of first impressions formed during early patient–doctor encounters in establishing a stable and positive therapeutic alliance (van Dulmen et al., 1997). Social attributions such as competence, warmth, and patient-centeredness can foster more favorable attitudes toward physicians, which in turn may enhance patient satisfaction and adherence to treatment plans (James, 2016). Furthermore, the evidence highlights the powerful impact of appearance- based cues in healthcare interactions. These cues not only influence providers' judgments and attitudes toward patients but also shape patients' perceptions of their doctors, ultimately affecting the quality of their relationship and relevant medical outcomes. Crucially, these effects are not limited to patients' satisfaction or adherence but extend to more fundamental dimensions of the clinical encounter- most prominently, the experience of pain.

1.2.3.1 The effect of social attributions on pain experience

Patients' facial-based social attributions from healthcare providers' appearance emerged as crucial factors in influencing their subjective experience of medical procedures- including pain (Bagnis et al., 2023; Mattarozzi et al., 2021; Necka et al., 2021).

Laboratory studies evidenced that participants expected less post- procedural pain and to be less likely to use prescription- strength pain medication following procedures conducted by providers they judged to be more competent, warm, and similar to themselves (Necka et al., 2021). Moreover, inferences of intrinsic familiarity – namely, the physical resemblance of unknown individuals to people characterizing one's environment (Zebrowitz et al., 2007) – appear to predict patients' judgment of pain intensity during a real medical procedure (Mattarozzi et al., 2021), with higher perceived familiarity predicting lower pain intensity and higher treatment satisfaction. These findings are corroborated by data collected from a more recent laboratory study (Bagnis et al., 2023), showing that inferences of higher perceived familiarity correlated with lower levels of pain intensity.

A possible interpretation of the effects that patients' social attributions from healthcare provider's facial appearance exert on their overall medical experience may be still rooted in the overgeneralization hypothesis (Zebrowitz et al., 2003). As previously mentioned, individuals are inclined to associate specific facial features, as well as the social attributes they convey, with particular social group memberships and occupations (Bin Meshar et al., 2022; Stolier et al., 2020). For instance, White, wealthy- looking males with prominent jawlines are more likely to be associated with the occupation of an "investment banker" than a "nurse", as their high perceived dominance aligns more strongly with the social attributes expected in the former than in the latter profession

(Oldmeadow et al., 2013). In the healthcare context, and particularly with regard to doctors, certain facial features linked to specific social attributes may be especially effective in shaping patients' experiences of medical procedures- especially when such features align with the social stereotype of a "good doctor". Among these features, in line with the SBT and the most recent findings on the topic (Bagnis et al., 2023; Mattarozzi et al., 2021), perceived familiarity may be particularly relevant. Indeed, high familiar- looking doctors may be especially reassuring, and hence triggering feelings of safeness that will ultimately reduce pain and increase satisfaction.

Collectively, these findings indicate that healthcare providers' facial appearance can influence the subjective experience of pain (Bagnis et al., 2023; Mattarozzi et al., 2021). Nonetheless, given the role of individual differences– including genetic factors–in shaping both pain perception and face evaluation, this line of research calls for a deeper exploration of how these differences contribute to the social modulation of pain induced by facial-based inferences of social judgments.

1.3 Individual Differences, Genetic variants and the social modulation of pain

The presence of conspecifics can influence the subjective experience of pain, potentially leading to placebo hypoalgesic effects (Che et al., 2018; Krahé et al., 2013). This phenomenon may have evolved as an adaptive mechanism, helping individuals cope with aversive experiences—such as pain—when others are present. Yet, not all individuals benefit equally.

Psychological and dispositional factors can shape the social modulation of pain, either enhancing or hindering its effectiveness. For instance, people high in pain catastrophizing– that is, the tendency to exaggerate the threat of pain and see themselves as unable to cope with pain (Keefe et al., 2000) – may exaggerate their pain expressions and report higher pain intensity (Sullivan, 2012) to attract support from others. Moreover, individuals with an anxious attachment style, that is characterized by fear of abandonment and need for reassurance from others (Mikulincer et al., 2009), reported higher pain ratings in the low- empathy– compared to the high-empathy– condition (Sambo et al., 2010) and benefited the most from verbal support (Jaltare et al., 2024). Differently, individuals with an avoidant attachment style, who usually trust less of others (Mikulincer & Shaver, 2010), reported higher pain in the social presence condition vs alone (Sambo et al., 2010).

However, this inter individual variability may depend not only on psychological traits, but also on more biological factors– as the genetic profile. Recent findings indeed suggest that genetic variation affects individual pain sensitivity (Colloca et al., 2019) as well as social- affective components such

as emotion regulation (Drabant et al., 2006), prosociality (Li et al., 2015), social responsiveness (Melchers et al., 2015, 2017) and stress resilience (Mayo et al., 2020; Serrano et al., 2019) – all aspects that intervene in the social modulation of pain. Specifically, research on these topics focused on the effects of genetic variations in four different neurotransmitter systems: the dopaminergic and the endocannabinoid systems – mostly studied in relation to placebo hypoalgesia (Colloca et al., 2019; Wrobel et al., 2014); and the serotoninergic and oxytocinergic systems– involved in both face processing and sociality. To enhance clarity, findings will be discussed separately for each neurotransmitter system.

Among the genetic polymorphisms involved in placebo hypoalgesia, the exonic single nucleotide polymorphism (SNP) rs4680, in the catechol-o-metyltransferase (COMT) gene, has been one of the most investigated. This SNP results in a valine (Val) to methionine (Met) substitution at codon 158 (Val158Met), and comprises two allelic variants: the G allele (coding for Val) and the A allele (coding for Met). Notably, the A allele is associated with a three- to four-fold reduction in COMT enzymatic activity (Lotta et al., 1995), leading Met (A) carriers to exhibit higher dopamine levels in the prefrontal cortex.

The neurochemical differences between the A and the G alleles have been principally linked to pain sensitivity and placebo responsiveness, with A carriers generally reporting higher pain sensitivity (Schmahl et al., 2012; Vargas-Alarcón et al., 2007) and enhanced placebo responses (Hall et al., 2012; Yu et al., 2014). However, this polymorphism received considerable attention also for its role in the processing of socially salient stimuli—such as faces—and in social-affective functioning in general. Individuals with the AA genotype have been shown to exhibit greater sensitivity to facial configuration changes (Doi et al., 2015) and better performance in facial emotion recognition task (Soeiro-de-Souza et al., 2012). Moreover, AA carriers showed increased vulnerability to emotional dysregulation (Barzman et al., 2015; Smolka et al., 2005) and impairment in disengaging attention from threat- related information (Herrmann et al., 2009). Differently, GG carriers have been associated with more effective emotion regulation (Drabant et al., 2006), higher levels of altruism (Reuter et al., 2011) and greater resilience to stress (Serrano et al., 2019).

To account for these divergent profiles, the COMT warrior/worrier model has been proposed (Goldman et al., 2006; Serrano et al., 2019). According to this framework, carriers of the A and G allele will react differently to stressful situations due to differences in tonic dopamine levels. Specifically, GG carriers– the "warriors" – will benefit from lower dopaminergic levels, which will enhance emotional stability and stress resilience. Conversely, AA carriers– the "worriers" – will

resent of higher prefrontal dopamine levels, exhibiting heightened affective reactivity (Kilford et al., 2015). In light of this model, it is possible that in stressful contexts—such as pain context—socially salient cues as faces may be differentially effective depending on genotype. While G carriers may be more likely to interpret such stimuli as "safety signals", thereby experiencing a buffering effect on pain, [1]A* carriers may engage with them less adaptively, perceiving them as irrelevant—or even threatening.

As with dopaminergic transmission, endocannabinoid signalling has also been implicated in placebo hypoalgesia (Benedetti et al., 2011). Within this framework, research focused on a specific functional variant in the Fatty Acid Amino Hydrolase (FAAH) gene, namely the SNP rs324420. This SNP encodes a Pro129Thr missense substitution (Cajanus et al., 2016), resulting in the A and the C variants. Notably, the A allele is associated with reduced FAAH enzymatic activity, which in turn leads to elevated levels of anandamide (AEA), an endogenous ligand of the CB1 receptor (Chiang et al., 2004; Sipe et al., 2002).

CC carriers exhibit stronger placebo hypoalgesic responses compared to [2]A* carriers (Peciña et al., 2014). Moreover, when considered together with COMT AA, homozygotes of this allelic form report a stronger placebo hypoalgesic response compared to carriers of other genetic variants (Colloca et al., 2019).

Concerning social-affective behaviours, instead, A* carriers seem to report decreased threat- related amygdala reactivity and reduced anxiety compared to CC carriers (Hariri et al., 2009; Mayo et al., 2020), as well as increased reward sensitivity (Hariri et al., 2009) and better emotion regulation (Gärtner et al., 2019).

Therefore, while CC carriers tend to report stronger placebo analgesia in response to placebo administration, FAAH A* carriers– characterized by enhanced emotion regulation, reduced anxiety, and increased sensitivity to rewards– may benefit the most from socially mediated pain modulation.

[1] Individuals carrying either the AG or the GG allelic form.

[2] Individuals carrying either the AA or the AC allelic form.

Although evidence linking the serotoninergic system to placebo analgesia is scarce (Diniz et al., 2018), this neurotransmitter system seems to be involved in both emotion regulation (Inoue et al., 2010) and sensitivity to social stimuli as faces (Brown et al., 2005). In this regard, particular attention has been given to the functional rs4570625 SNP in the TPH2 gene, which encodes the enzyme tryptophan hydroxylase 2—crucial for serotonin synthesis. TPH2 rs4570625 determines a G to T base substitution in the promoter region at position -703, resulting in two different variants: the G and the T variants.

Concerning placebo response, GG carriers reported placebo- induced reductions in both anxiety symptoms and stress- related activity in the amygdala (Furmark et al., 2008). In terms of social-affective behaviours, instead, data are quite mixed. If some studies evidenced [3]T* carriers as more likely to show emotion dysregulation (Inoue et al., 2010) and affective disorders (Gao et al., 2012), more recent findings suggest how homozygotes for the T variant may report lower levels of aggressiveness, anxiety and depressiveness (Laas et al., 2017). Additionally, animal studies suggest that T* carriers may be more environment- sensitive compared to GG carriers, while human studies evidenced a stronger amygdala response when positively and negatively valenced faces were presented (Brown et al., 2005).

Taken together, the current findings on TPH2 rs4570625 suggest that the T* allele of this SNP may be particularly relevant when socially salient stimuli—such as faces—are involved. Therefore, despite the limited and mixed information, this genetic variant may still represent a promising candidate for investigating the genetic bases of social modulation of pain.

Finally, since the pivotal role oxytocin plays in social cognition (Ebert & Brüne, 2018), genetic variations linked to this neurotransmitter system have been extensively studied in relation to social, prosocial and affective behaviours. Specifically, the SNP rs2254298 determines a G to A base substitution in the intron 3 region of the OXTR gene, resulting in the A and the G variants of this gene. A carriers show greater amygdalar volumes bilaterally compared to G carriers, which may cause increased risk for developing autism (LoParo & Waldman, 2015), depression and anxiety disorders (Brüne, 2012). Moreover, [4]A* carriers seem to recognize emotional facial expression less frequently

[3] Individuals carrying either the TT or the TG allelic form

[4] Individuals carrying either the AA or the AG allelic form

than GG carriers (Lucht et al., 2013), and show greater social impairment during early age (Parker et al., 2014).

It derives that individual differences in OXTR polymorphisms appear to influence sensitivity to social stimuli and prosocial tendencies. Genetic variations of this gene may therefore offer a plausible basis to investigate more deeply how the presentation of social salient stimuli, as faces, differently influence individuals' perception of pain.

Collectively, these findings corroborate the hypothesis that genetic variability may actually play a significant role in shaping the social modulation of pain. Individuals carrying genetic variants associated with increased prosociality and social sensitivity—such as those enhancing face processing and affiliative behavior—alongside variants linked to lower pain sensitivity and stronger placebo responsiveness, may be especially likely to benefit from social support; even when simply primed by the presentation of social salient stimuli as faces. The consequence of this modulatory mechanism is twofold. First, it reinforces the idea that the social modulation of pain may have evolved as an adaptive mechanism to increase survival in the presence of conspecifics. Second, it underscores the powerful role of human faces as social salient stimuli capable of altering pain perception. Faces do not merely signal the presence of another person: they inform about the surrounding context, the emotion perceived by the expresser but also the expresser themselves. Yet, not all faces are perceived equally. Depending on the social attributions they evoke, a face may either enhance or reduce the perceived intensity of pain. This effect becomes especially relevant in medical contexts, where the face of a healthcare provider can influence not only the therapeutic alliance, but also relevant medical outcomes as perceived pain.

In this regard, facial features conveying specific social attributes– and more specifically, those aligning with the stereotypical representation of a "good doctor" – may be especially relevant in determining this modulatory effect. Among these traits, familiarity may be particularly influential, both in light of the existing evidence and of its possible evolutionary function in fostering cooperation and enhancing survival (Behrmann & Avidan, 2022).

Together, these considerations highlight the need to systematically identify the physical features and social attributes defining the stereotypical representation of a "good doctor," and to examine whether and how such representations, moderated by individual differences, influence the subjective experience of pain.

2 Research projects and Objectives

Across four different studies, the present PhD project seeks to bridge two lines of investigation– namely the social modulation of pain and the impact of facial-appearance based inferences on behaviour– aiming to pursue four different and interrelated goals.

The first study focused on developing and validating the mental and stereotypical representation of a "good doctor" at the group level. Using ratings from an Italian sample and machine- learning methods (Peterson et al., 2022), we aimed to identify the physical facial features– and, most importantly, the associated social traits– that define this representation. Building on previous findings (Bagnis et al., 2023; Mattarozzi et al., 2021), and in line with the overgeneralization hypothesis (Zebrowitz et al., 2003, 2007; Zebrowitz & Montepare, 2006), we expected to find that the social dimension of familiarity would play a crucial role in characterizing the "good doctor" stereotype. Furthermore, we investigated whether this representation may be affected by the motivation to become part of this social category, and hence if it may differ in a group of medical students from the general population.

The second study, instead, investigated whether this stereotypical representation modulates the subjective pain experience by reducing the sense of threat through inferences of familiarity. In accordance with the SBT (Coan & Sbarra, 2015), we predicted that pain perception would decrease when painful stimulation was concurrently administered with the processing of a face representing a "good doctor". Moreover, in light of the latest findings highlighting the pivotal role of genetics in shaping both pain modulation and socio-affiliative behaviours, we also investigated how individual differences in genetic profiles contribute to this modulatory effect. Consistent with the literature, we hypothesized that individuals carrying genetic variants related to better emotion regulation (Drabant et al., 2006), greater stress resilience (Kilford et al., 2015) and reduced anxiety in response to stressors (Mayo et al., 2020; Spagnolo et al., 2016), as well as heightened responsiveness to social stimuli (Brown et al., 2005) and high levels of prosociality (Brüne, 2012), would benefit the most from the presentation of faces reflecting the stereotypical representation of a good doctor during painful stimulations, reporting lower levels of pain intensity and unpleasantness.

Due to the relevance of idiosyncratic differences in forming complex social judgments in face evaluation (Albohn et al., 2025), the third study focused on creating and validating idiosyncratic (i.e., individual-specific) representations of the "good doctor" using generative models. This phase intends to corroborate the findings collected during the first part of the project; highlighting the individual components of the "good doctor" stereotype and emphasizing perceived familiarity as a central trait.

Finally, the fourth study aimed to evidence whether these idiosyncratic representations of the "good doctor" stereotype could be more effective in modulating pain perception compared to non-idiosyncratic representations.

Together, this PhD project aims to provide a comprehensive framework for understanding how facial cues – particularly those that evoke the mental and stereotypical representations of a "good doctor", giving rise to perceptions of familiarity and social relevance– serve as safety signals capable of activating endogenous pain-inhibitory mechanisms, akin to placebo hypoalgesia. By bridging the fields of social cognition and pain modulation, it seeks not only to advance theoretical models of placebo mechanisms but also to inform the development of socially grounded strategies in clinical contexts.

3 A data-driven method for visualizing the "Good doctor" stereotypical representation

3.1 Introduction

Healthcare providers' facial appearance have shown to exert a profound impact on relevant clinical outcomes, such as pain perception (Bagnis et al., 2023; Mattarozzi et al., 2021), post-procedural medication use (Necka et al., 2021) and treatment satisfaction (Mattarozzi et al., 2021). However, these studies have mainly focused on the effects that inferences of specific social traits – such as familiarity and competence for instance (Bagnis et al., 2023; Mattarozzi et al., 2021; Necka et al., 2021) – have in shaping patients' experiences, without further investigating which physical facial features- and the associated personality traits- people expect to perceive in a doctor to be judged as "good" at first glance.

In light of this gap, the present study aims to identify the stereotypical representation of a "good doctor" by capitalizing on computational, data-driven methods. These methods allow to visualize the stereotypical mental representations of complex social attributes and to model them through systematic variations in facial features (Todorov et al., 2025; Todorov & Oh, 2021). Moreover, since models on other social attributes have already been computed within the same multidimensional face space used to generate faces (Todorov et al., 2025), the similarity between the new "good doctor" model and these models can be readily assessed, thereby identifying which social attributes and physical traits define this stereotypical representation. Here, we build on recent research that used hyper-realistic faces to model the representations of 35 socially meaningful social attributes, such as perceived competence, dominance, familiarity, and trustworthiness (Peterson et al., 2022). Using the same method, we constructed a new model dimension of the "good doctor" mental representation, which can be applied to manipulate novel facial stimuli by systematically increasing or decreasing their similarity to this representation.

We conducted two experiments to a) collect behavioral data as an input to the model (Experiment 1) and b) to validate the resulting model (Experiment 2). Specifically, in Experiment 1 participants rated randomly sampled hyper- realistic facial images on their resemblance to a "good doctor." Using these ratings as input for predictive models in the latent face space, we created a new "good doctor" dimension, which captures the physical features—as well as the related social attributes—defining this stereotypical representation. Drawing on previous findings (Bagnis et al., 2023; Mattarozzi et al.,

2021) and in line with the overgeneralization hypothesis (Cuddy et al., 2008; Zebrowitz et al., 2003, 2007; Zebrowitz & Montepare, 2006), we expected to find a strong correlation between the "good doctor" dimension and the familiarity attribute (for more details, see Peterson et al., 2022).

In Experiment 2, facial stimuli were manipulated to vary in their resemblance to the defined "good doctor" dimension. An independent sample was recruited to evaluate them based on the likelihood of recommending them as a doctor to others. Aligned with predictions of the "good doctor" model dimension, we expected that "high good doctor" faces, namely faces manipulated to closely represent a "good doctor", would be more likely to receive recommendations compared to "low good doctor" faces, which deviate from that stereotypical mental representation.

3.2 Experiment 1

3.2.1 Materials and Methods

3.2.1.1 Participants

393 participants (261 women, 130 men,2 non-binary, age range: 18-71 years, *M*=27.90; *SD*=11.90 years) were recruited via e-mail and during lectures. Participants self- identified as follows: 381 Caucasian, one Balkan, six Latin American, three Afro American, one Hispanic and one Asian.

Participants represented three different subsamples, namely general population (*N*=166, 87 women, 78 men, one non-binary, age range: 18-71 years, *M*=36.42, *SD*=13.72 years), second-year medical students enrolled in the Medical course of the University of Bologna (*N*=74, 46 women, 28 men, age range: 19-31; *M*=20.78, *SD*=1.93 years) and students enrolled in other healthcare courses (i.e., Nursing or Dental Hygienist courses) of the same University (*N*=153, 128 women, 24 men, one non-binary, age range: 19- 43; *M*=22.11; *SD*=4.66). Notably, students enrolled in other healthcare courses are convenience participants at the very beginning of their educational paths, having not yet completed their mandatory hospital training or clinical internships. Their stereotypical model of the "good doctor" is therefore unlikely to be influenced by practical clinical experience at this stage. However, their motivation to pursue a healthcare career was considered in the subsequent analyses (see subchapter 3.2.2). Of this initial sample, 25 participants (12 from the "General population" sample, 3 from the "Medical students" sample and 10 from the "other healthcare students" sample) were excluded either for scarce compliance to the task (N=15) or because they showed test- retest reliabilities at or below 0 (N=10). Therefore, the final sample consisted of 368 participants (247 women, 119 males,2 non binary; age range: 18-71 years, *M*=27.83; *SD*=11.88).

Sample size fits within the suggested sample size needed to achieve a stable mean trait impression rating from faces (i.e., around 30 rates for each personality trait) (Hehman et al., 2018; Peterson et al., 2022).

The study was reviewed and approved by the Bioethics Committee of the University of Bologna (Prot. No. 0394328).

3.2.1.2 Stimuli

Facial stimuli were selected from the pool of 1004 synthetic face images used by Peterson et al., 2022 to develop their One Million Impression Dataset (all the synthesized images and the associated behavioral judgments are deposited in a GitHub repository at the following link: https://github.com/jcpeterson/omi.). Children and adolescents' faces (under 18 years old) were excluded. Hence, a total number of 849 pictures of both male and female faces were selected for task construction.

Each facial stimulus was rated using a 0 (Not at all) to 100 (Extremely) Visual Analogic Scale (VAS) by an independent sample of 4,157 participants from a previous study (Peterson et al., 2022) along the following 35 physical features and social attributes: "gender (masculine/feminine)", "age", "weight", "long- haired", "hair-color", "skin-color", "white", "black", "asian", "middle- eastern", "hispanic", "islander", "native", "happy", "familiar", "looks-like-you", "typical", "trustworthy", "attractive", "outgoing", "outdoors", "healthy", "well- groomed", "dominant", "cute", "alert", "smart", "electable", "smug", "dorky", "privileged", "liberal", "memorable", "gay", "godly/religious".

The selected faces were distributed across three different links via Qualtrics to ensure task feasibility (i.e., less than 30 minutes long). Each link presented a single version of the task, composed of 283 facial stimuli. To assess intra-rater test-retest reliability, 20% of the administered stimuli were presented twice (Link 1: 56 stimuli; Link 2: 58 stimuli; Link 3: 56 stimuli). This resulted in approximately 349 images for each version of the task.

3.2.1.3 General procedure

Participants were invited to participate in a study on «the effects of first impressions on individual's behavior and decisions in the healthcare field». Once participants gave their consent, they were sent one of the three links via email. Hence, each participant completed a single version of the task.

Participants were instructed to indicate, on a VAS from 0 (“Not at all”) to 100 (“Extremely”), how much the to-be presented faces resembled that of a “good doctor” according to their first impression. Two practice trials were then administered to become familiar with the procedure. Then, the selected 283 facial stimuli were randomly presented one after the other, followed by the question “To what extent do you think this person looks like a “good doctor”?”.

An example trial is illustrated in Figure 3.1.

Figure 3.1

Example trial of the task administered during Experiment 1

To minimize distraction, participants were allowed to pause while completing the task. The task lasted about 30 minutes.

Participants were fully debriefed about the research question at the end of their participation.

3.2.1.4 Data analyses

To assess test- retest reliability for each participant, we computed the correlation between the ratings participants gave to repeated stimuli. Participants with correlations less than or equal to zero ($r \leq 0$) were excluded from further analyses (Peterson et al., 2022). Additionally, to assess the internal consistency of participants’ ratings, we computed a split-half correlation coefficient by dividing each participant’s ratings into two halves across the presented faces and calculating the correlation between the mean ratings of each half.

To explore the structure of social inferences underlying the “good doctor” attribute, we first conducted a Principal Component Analysis (PCA) on the 35 social and psychological attributes from

the One Million Impressions Dataset (Peterson et al., 2022). For each facial stimulus, mean ratings on each attribute were computed and entered into the analysis. Components were extracted using Principal Components Analysis with Varimax rotation, and sampling adequacy was assessed using the Kaiser–Meyer–Olkin (KMO) measure and Bartlett's test of sphericity. Moreover, cases with missing data were excluded via listwise deletion. The Kaiser criterion (eigenvalues >1) was applied to determine the number of components to retain, and only loadings with an absolute value greater than 0.30 were interpreted.

To examine the contribution of these latent social dimensions to perceptions of a "good doctor," we subsequently conducted a multiple regression analysis. Factor scores for each extracted PCA component were computed for each facial stimulus, and were entered in the regression analysis as predictors of the mean "good doctor" ratings.

To generate the "good doctor" model, we computed the mean ratings of participants for each presented synthetic face. Each face was represented as a 512-dimensional latent vector obtained from the StyleGAN2 generator. We then fit a separate L2-regularized linear regression (ridge regression) mapping the latent vectors $z_i \in R^{512}$ to the average ratings y_i the "good doctor" attribute:

$$y_i = w_0 + w_1 z_1 + ... + w_d z_{id} + \epsilon$$

where w_0 is the intercept, z_{ij} is the j-th component of the latent vector of the i-th face, w_j is the weight of that component in defining the "good doctor" rating and ϵ is the residual error. The resulting coefficient vector

$$w=\{w_1,\ldots,w_d\}$$

defines the "good doctor" dimension in the latent space.

This weight vector can be used both to predict perceived "good doctor" ratings of new faces and to manipulate face representations along this dimension via vector arithmetic:

$$z_i^* = z_i + \beta_w$$

where β controls the strength and direction of the manipulation.

Finally, we computed pairwise correlations between the PCA components and the "good doctor" mean ratings for each facial stimulus. We also extracted the weight vectors of all attributes from an omnibus model and computed pairwise correlations between these vectors and the "good doctor"

weight vector. This correlation matrix quantifies the degree of overlap or shared representational structure between the “good doctor” construct and other social attributes (Peterson et al., 2022).

Analyses were performed using SPSS (version 20.0; Chicago, IL).

3.2.2 Results

The three subsamples considered (i.e., general population, medical students and other healthcare students) differ in terms of sex (χ^2 (2, N=368) = 32.71; p<.001) and age (F(2,365)=110.79; p<.001). The split-half correlation coefficient for each group (general population: r=.80; medical students: r=.69; other healthcare students: r=.79) showed that participants' ratings reliably reflect the degree to which the displayed face represented, in their opinion, a "good doctor”.

Results of the KMO measure and the Bartlett’s test of sphericity indicated that the data were suitable for PCA (KMO=.87, Bartlett’s test of sphericity: $\chi^2(595)=35565.36$, p< .001). PCA yielded a four-components solution, explaining the 64.85% of the total variance after Varimax rotation.

Component 1, explaining the 23.78% of the variance, was primarily loaded by ethnicity-related (e.g., “White”, “Islander”, “Native”, “Hispanic”) and perceived similarity (e.g., “privileged”, “looks like you”, “typical”). Notably, the White and perceived similarity attributes loaded in the opposite direction relative to the other ethnic attributes. Component 2, explaining the 21.89% of the variance, reflected attributes related to attractiveness (e.g., “attractive”, “cute”, “healthy”, “age”): positive loadings were associated to attractive- related features as “attractive”, “cute” and “healthy”; whereas negative loadings corresponded to attributes as “age”, “gender” (masculine vs feminine, reflected by lower and higher ratings on a 0-100 VAS, respectively) and “weight”. Component 3, accounting for the 10.46% of the variance, grouped attributes related to positive emotional expressions (e.g., “happy”, “outgoing”) and trustworthiness, which loaded positively, and dominance, which loaded negatively. Finally, Component 4, explaining the 8.72% of the variance, reflected attributes related to familiarity (e.g., “familiarity”), professional look (e.g., “electable”) and intelligence (e.g., “smart”).

The rotated component matrix for the four components solution is presented in Table 3.1. Loadings ≥ |0.30| are reported.

Table 3.1

Rotated component matrix for the four components solution

	Components			
	1	**2**	**3**	**4**
	(Ethnicity and perceived similarity)	**(Attractiveness)**	**(Trustworthiness)**	**(Familiarity and professional look/intelligence)**
White	**-.94**			
Islander	**.90**			
Native	**.85**			
Skin color	**.83**			
Middle eastern	**.79**			
Hispanic	**.78**			
Black	**.71**			
Asian	**.70**			
Hair color	**.66**			
Privileged	**-.65**	.30		.41
Looks like you	**-.60**			
Typical	**-.57**			.57
Liberal	**.49**	.38	.37	-.46
Attractive		**.90**		
Cute		**.87**	.41	
Healthy		**.80**	.31	
Age		**-.74**		.37
Long haired		**.74**		
Gender		**-.68**	-.42	
Dorky		**-.67**		-.43
Well groomed		**.65**		
Memorable	.31	**.53**		
Weight		**-.53**		
Happy			**.92**	

Trustworthy			**.88**	
Outgoing			**.85**	
Smug			**-.81**	
Dominant		-.34	**-.77**	.35
Alert			**.68**	
Electable				**.83**
Familiar	-.41			**.64**
Godly			.50	**.63**
Gay		-.33		**-.58**
Smart			.33	**.54**
Outdoors				**-.32**

Note. Loadings ≥ |0.30| are shown. The highest factor loadings for each variable are evidenced in bold.

The multiple regression analysis using the four PCA components as predictors of the "good doctor" attribute yielded a significant model ($F(4,844)=440.78$, $p<.001$), explaining 68% of the variance ($R2=.68$). Ethnicity and perceived similarity, as well as attractiveness, showed a significant negative association with the "good doctor" attribute (Ethnicity and perceived similarity: $\beta = -.17$, $p<.001$; attractiveness: $\beta = -.12$, $p<.001$). Conversely, trustworthiness, as well as familiarity and professional look/intelligence, showed significant positive associations (trustworthiness: $\beta = .22$, $p<.001$; familiarity and professional look/intelligence: $\beta = .76$, $p<.001$). Among all predictors, familiarity and professional look/intelligence emerged as the strongest contributor to judgments of the "good doctor" attribute.

A "good doctor" model dimension was then generated for the entire sample and separately for each of the three subsamples considered (i.e., general population, medical students and other healthcare students), resulting in 4 distinct "good doctor" models. The four "good doctor" models showed good fit to the data ($R2=.50$) and were highly correlated with one another ($rs >.90$). Given this high degree of convergence, we decided to refer exclusively to the 'good doctor' model derived from the full sample.

Transformation of the facial stimuli along the "good doctor" dimensions is shown in Figure 3.2. The central picture of each row represents faces in their "neutral" version, which is how they were presented to participants. On the extreme left and right, the same stimulus is displayed when

manipulated to -3 and +3 SD from the mean, respectively. Stimuli at +3 SD represent the “high good doctor” version of the facial stimuli; featuring facial features that, according to the sample considered, are typically associated with the stereotype of a good doctor. Differently, stimuli at -3 SD represent the "low good doctor" version of these stimuli, presenting facial features distant from those of a “good doctor”.

Figure 3.2

Model- based manipulations along the “good doctor” dimension.

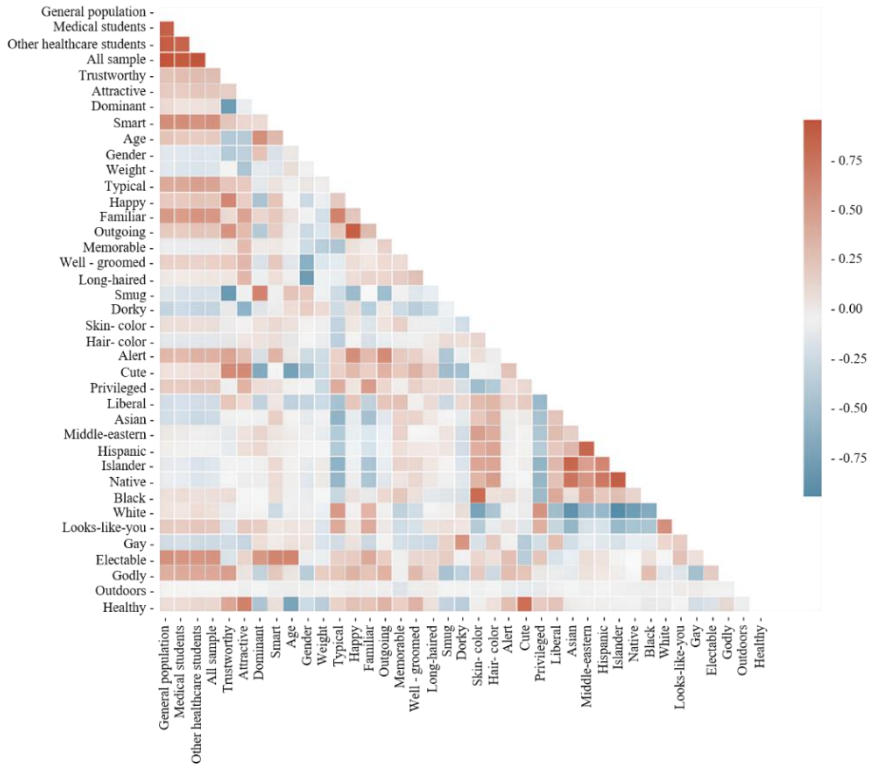
-3 SD 0 +3 SD

Finally, correlations between this model and the 35 social and psychological attribute models included in the One Million Impression Dataset (Peterson et al., 2022) mirrored the regression results. Indeed, weak correlations emerged between the “good doctor” model dimension and the “white” (r=.11), “looks like you” (r=.27), “attractive” (r=.26) and “gender” (r=-.10) attribute models, whereas stronger correlations were found with the “trustworthy” (r=.33), familiar (r=.58), “electable” (r=.60) and “smart” (r=.61) attribute models.

The correlation matrix between the “good doctor” model dimension and the 35 social and psychological attribute models included in the One Million Impression Dataset (Peterson et al., 2022) is displayed in Figure 3.3.

Figure 3.3

Correlations between the "good doctor" dimensions and the 35 attributes of the One Million Impression Dataset (Peterson et al., 2022).



3.3 Experiment 2

Experiment 1 focused on developing the "good doctor" model dimension, which allows the manipulation of facial images to assume the physical features characterizing the stereotypical representation of a "good doctor" in a sample of Italian participants. However, empirical validation is still required to assess whether manipulations along this new dimension can influence socially relevant judgments. In particular, we aimed to measure decisions that are modulated by this stereotypical representation, in order to evaluate its influence on social attributions and preference within a decision-making context. In Experiment 2, we therefore manipulated novel facial stimuli along the "good doctor" dimension and asked an independent group of participants to evaluate them. This procedure allowed us to test the predictive validity of the model, examining whether faces resembling the "good doctor" stereotype would be more likely to be recommended in a medical context compared to those deviating from it.

3.3.1 Materials and Methods

3.3.1.1 Participants

257 Caucasian participants (179 females, 78 males, age range: 18 -84 years, *M*=32.48; *SD* = 15.41 years) were recruited by fliers, via e-mail and during lectures. 14 participants were further excluded because they showed low intra-rater reliability ($r < 0$). Therefore, the final sample consisted of 242 participants (170 females, 74 males, age range: 18-84 years, *M*=32.16; *SD*=15.29 years).

Sample size was established through an a- priori calculation using G*Power software (Faul et al., 2007), which indicated that assuming a medium effect size (*f*=0.25) and a number of three conditions (i.e. high good doctor/low good doctor/neutral faces), a minimum of 159 would be needed to achieve a statistical power of 0.8 for alpha=0.05.

The experimental procedure was approved by the Institutional Review Board of the University of Bologna, Italy (Prot. No. 0394328).

3.3.1.2 Stimuli

All Caucasian identities (*N*=183) of the Chicago Face Database (Ma et al., 2015) have been manipulated along the previously identified “good doctor” dimension (i.e., the “good doctor” dimension generated considering the sample as a whole). Manipulations occurred encoding the faces from the CFD into StyleGAN2's latent space, resulting in a 512-dimensional vector for each face. These vectors were then multiplied by the previously generated “good doctor” dimension and a scalar (i.e., +3 and -3), resulting in three different versions for each identity: the "high good doctor" (+3SD from the mean of the “good doctor” dimension), the "low good doctor" (-3SD from the mean “good doctor” dimension) and the neutral version (the stimulus in its original version). Therefore, a total of 549 stimuli were used in the study.

Task administration occurred through four different links via Qualtrics. Once again, we divided the stimuli into four links, each designed to take less than 15 minutes to complete. Each link contained a single version of the task, which was composed of 137 stimuli (except for Link 1, which contained 138 stimuli) presented in a fully randomized order. For test-retest reliability, 20% of the administered stimuli (i.e., 27 stimuli) were presented twice, for a total of about 164 stimuli for each version of the task.

3.3.1.3 General procedure

Participants were invited to participate in a study on "the effect of facial stimuli perception on individual's behavior and decisions in the healthcare field". After providing informed consent, they received one of the four links during lectures or via e-mail. Therefore, each participant completed a single version of the task.

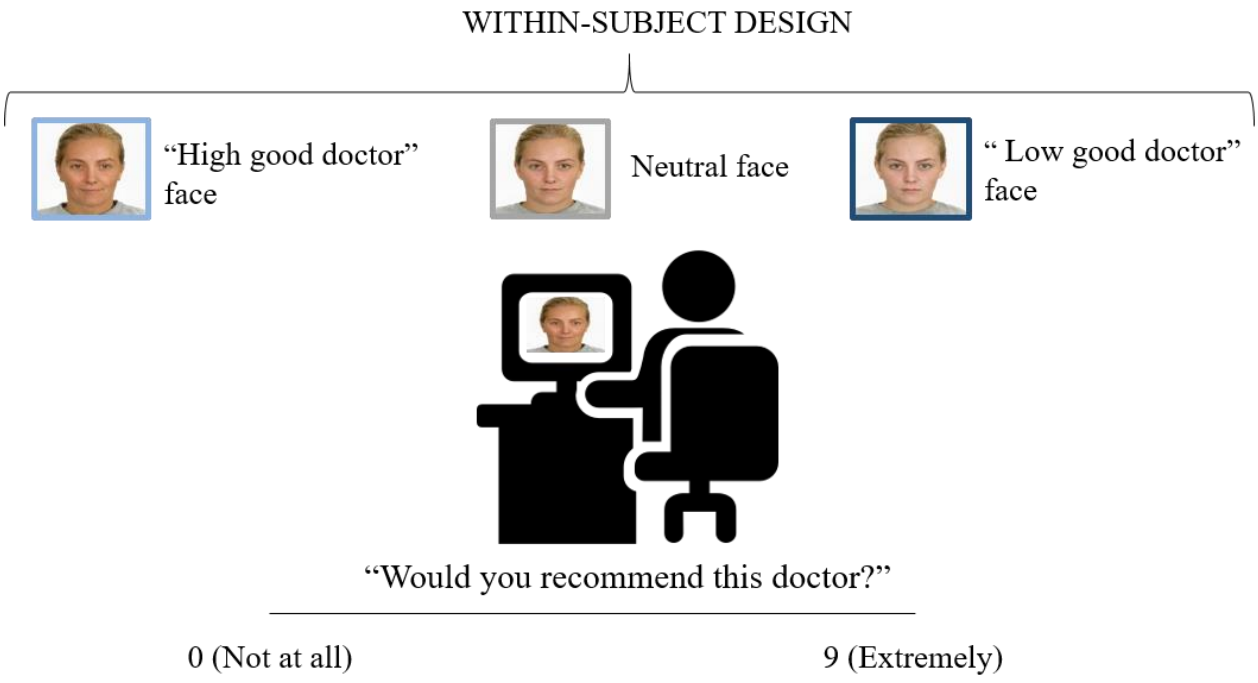
Participants were instructed to look at the facial stimuli representing different doctors and to respond based on their first impression. Before starting, two practice trials were administered to familiarize them with the procedure. Then, facial stimuli were randomly presented one after the other.

For each picture, participants were asked to indicate, on a visual-analogic scale (VAS) from 0 ("Not at all") to 9 ("Very much so"), how likely they were to recommend the displayed doctor. To minimize distraction, they were allowed to pause while completing the task.

An example trial is illustrated in Figure 3.4.

Figure 3.4

Example trial of the task administered during Experiment 2



Participants were fully debriefed about the research question at the end of their participation.

3.3.1.4 Data analyses

To assess test- retest reliability, we calculated intra- rater correlations between the scoring of the repeated stimuli the first vs. the second time they were presented for each participant. Participants who had negative or equal to zero ($r \leq .0$) intra- rater correlation were excluded from the analysis (Peterson et al., 2022).

A one-way analysis of variance (ANOVA) was run to evidence whether significant differences could be observed between the means of the ratings participants gave in the three different conditions. The likelihood to recommend the displayed doctor was considered as the dependent variable, while the "good doctor" dimension ("high good doctor", "low good doctor", "neutral") was considered as the independent variable. Effect sizes were calculated using partial eta squared.

The alpha level for all analyses was set to $p<.05$.

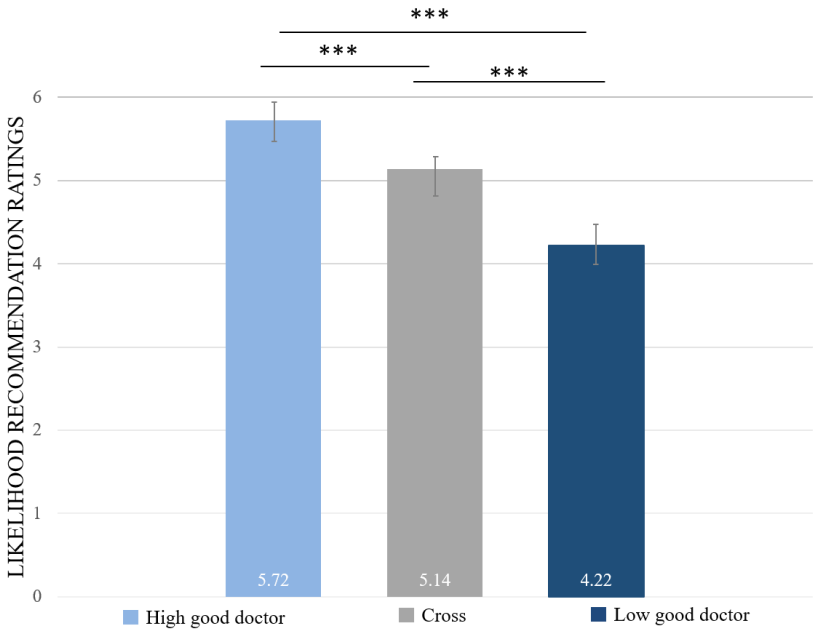
Analyses were performed using SPSS (version 20.0; Chicago, IL).

3.3.2 Results

As shown in Figure 3.5, significant differences were found between "high good doctor", "low good doctor" and "neutral" faces when the likelihood to recommend the doctor was considered ($F(2,546)=119.93$, $p<.001$, $\eta^p_2=.31$). Specifically, "high good doctor" faces (M=5.72; SD=0.92) were significantly more likely to be recommended by our participants compared to both "low good doctor" (M=4.22; SD=0.93) and "neutral" faces (M=5.14; SD=0.95). Moreover, a significant difference emerged between "neutral" and "low good doctor" faces, with the "low good doctor" faces being associated with the lowest probability of being recommended to others.

Figure 3.5

Mean differences in the likelihood of recommendation ratings for "high good doctor" faces, "low good doctor" faces and "neutral" faces.



Note. Error bars represent the standard error of the means. ****p*<.001

3.4 Discussion

Current research on doctors' first impressions has typically focused on single aspects of doctor's appearance, such as posture (Grün et al., 2022) and dress style (Rehman et al., 2005), or on psychosocial features, such as ethnicity and sex (Hall et al., 2020), attractiveness (Zhang et al., 2024) and familiarity (Bagnis et al., 2023; Mattarozzi et al., 2021). Here, we introduce for the first time a comprehensive framework—the "good doctor" model dimension—derived from responses across three distinct subsamples (general population, medical students, and other healthcare students). This dimension integrates both physical features and social attributes, providing a more holistic account of the stereotypical representation of a "good doctor". Moreover, this approach offers a novel model for understanding how physical features and social attributes jointly shape the perception of healthcare providers.

Consistent with the literature on the topic (Hall et al., 2020; Necka et al., 2021), Experiment 1 evidenced that judgments on the "good doctor" attribute were primarily influenced by social attributes

loading into four main social components- ethnicity and perceived similarity; attractiveness; trustworthiness; professional look/intelligence and, in line with our expectations, familiarity.

Physician's ethnicity and perceived similarity contribute to shaping perceptions of what a "good doctor" looks like. Patient-provider racial concordance— that is, the alignment between the racial identities of the patient and their healthcare provider—has been identified as a key factor in fostering an effective doctor-patient alliance (Shen et al., 2018). Patients in race-concordant dyads are more satisfied with the treatments received (Ku & Vichare, 2023a; Takeshita et al., 2020). Moreover, when given the option, patients tend to prefer doctors of their same race compared to other-race doctors (Ku & Vichare, 2023b; Simon & Lewis-Liu, 2024), reporting more positive outcomes in terms of partnership building, visit length, supportiveness and respectfulness (Shen et al., 2018). This preference for race-concordant doctors may reflect patients' preference for clinical interactions characterized by reduced racial bias (Saha & Beach, 2020), and alignment in cultural and communication style (Street et al., 2008). Moreover, as suggested by the loadings of perceived similarity attributes such as "typical" and "looks like you", it may also reflect perceived similarity favoritism and negative stereotyping toward other-race individuals (Bagnis et al., 2020; Zebrowitz et al., 2007). People from the same ethnicity are indeed perceived as more similar (Zebrowitz et al., 2007) and, in turn, more likeable (Montoya et al., 2008, 2017) and competent (Zebrowitz et al., 2008). Therefore, in relationships as personal as the doctor-patient relationship, patients may attribute substantial relevance to the physician's race, preferring race-concordant doctors (Saha & Beach, 2020). Accordingly, our results evidenced a covariation of perceived-similarity attributes with White-related attributes, consistent with the prevalence of White participants in the study.

Albeit modestly, attractiveness also contributed to defining the "good doctor" stereotype. Specifically, the negative regression coefficient associated with this component suggests that faces scoring higher on appearance-related cues typically associated with attractiveness and femininity were less strongly aligned with the "good doctor" stereotype. These results partially replicate previous findings on the topic: in a digital consultation context, for instance, physicians' facial attractiveness was negatively correlated with the number of visits to their homepage by users (Zhang et al., 2024b), although this effect was affected by gender-specialty bias. Moreover, patients showed a stronger preference for more attractive doctors of the opposite sex, whereas this preference attenuated for doctors of the same sex (Wei et al., 2024).

These findings may also be understood as reflecting the so- called "beauty is beastly" effect (Paustian-Underdahl & Walker, 2016), namely the tendency for attractive women applying for masculine sex-

typed jobs to be disadvantaged and discriminated (Johnson et al., 2014). Consistently, female doctors report that they still face difficulties in being fully recognized by patients in their professional role (Haynes et al., 2021; Stavely et al., 2022), despite the success in medical training and clinical practice women doctors have achieved (Tsugawa et al., 2017; Wallis et al., 2017) and their increasing representation in the physician workforce (Butkus et al., 2018). Together, these results highlight how the "good doctor" stereotype may be shaped by gendered social expectations, which associate this professional role with traditionally masculine attributes and, ultimately, with men.

Positive emotional expressions (e.g., happiness), and trustworthiness also contributed to build the image of a "good doctor". Happiness closely relates to perceptions of affiliation and trustworthiness (Oosterhof & Todorov, 2008), whereas trustworthiness itself emerged as a crucial factor in fostering a good doctor- patient relationship (Murray & McCrone, 2015) and in promoting better medical outcomes (Becker & Roblin, 2008; Birkhäuer et al., 2017).

The importance that trustworthiness seems to have in defining the "good doctor" stereotype is not surprising. Together with dominance, trustoworthness is indeed a pivotal dimension in face evaluation (Oosterhof & Todorov, 2008), signaling whether the other person's intent is for good or ill. Specifically, inferences of trustworthiness– i.e., the closest approximation of the valence of a face (Oosterhof & Todorov, 2008)– promote approach-oriented responses (Oosterhof & Todorov, 2008) and perceived warmth (Fiske et al., 2007); whereas dominance elicits feelings of potential threat. In healthcare, this may translate into patients' preference for healthcare providers who appear more trustworthy, warmer and caring; thereby shaping the "good doctor" stereotypical representation as embodying such characteristics.

Finally, familiarity and perceived professional look/intelligence emerged as the most influential components shaping the "good doctor" stereotype.

Familiarity has indeed been shown to be particularly relevant in the healthcare context. It correlates with lower self-reported pain (Bagnis et al., 2023; Mattarozzi et al., 2021) and greater treatment satisfaction (Mattarozzi et al., 2021). These effects resonate with the "familiar face overgeneralization" hypothesis, which posits that individuals often respond positively to new faces resembling familiar ones from their environment (Zebrowitz & Montepare, 2008). Such intrinsic familiarity triggers positive emotions (Dotsch et al., 2016; Zebrowitz et al., 2007), leading to feelings of safety and increased willingness to cooperate (Keller & Reeve, 1998; Vella et al., 2017). In

healthcare, this may translate into reduced patients' feelings of threat, fear, and anxiety, ultimately improving clinical outcomes (Bagnis et al., 2023; Mattarozzi et al., 2021).

Perceived professional look/intelligence appears equally relevant. Social attributes as “smart” and “electable” emerged as the most influential in shaping the “good doctor” stereotype. Specifically, smartness reflects people’s tendency to socially position themselves and others based on perceived efficacy, ability, and success (Sternberg, 2007; Hatt, 2012). Electability, instead, refers to the belief that an individual would prevail in a competitive setting (Minozzi & Woon, 2023), and encompasses various aspects of personality, with competence being among the most relevant (Ballew & Todorov, 2007; Chemers, 2014; Hall et al., 2009; Mosier & Pietri, 2021; Olivola & Todorov, 2010; Stone & Rapoport, 1994).

In this regard, competence has been often investigated in relation to doctor’s appearance (Ashton-James et al., 2019; Howe et al., 2019; Necka et al., 2021). Together with warmth, competence represents a universal dimension of human social cognition (Fiske et al., 2007), allowing individuals to differentiate each other in terms of perceived intent (i.e., warmth, capturing traits such as friendliness, helpfulness, and trustworthiness) and respect (i.e., competence, including intelligence, ability, and resource deployment). Furthermore, it reflects the possessor’s chances of achieving personal goals (Chemers, 2014). In relation to physicians’ appearance, experimental evidence indicates that individuals tend to prefer physicians who appear more competent (Necka et al., 2021). Moreover, higher perceived competence has been shown to predict greater trust in the physician and reduced pain during surgical procedures (Ashton-James et al., 2019), as well as lower expectations of pain both during and after the procedure (Necka et al., 2021). Therefore, it is likely that patients may prefer more professional, competent-looking doctors, which may enhance trust (Ashton-James et al., 2019) and reduce anxiety (Tedeschi et al., 2015).

It is noteworthy that the stereotypical representation of a “good doctor” does not seem to be influenced by motivation to belong to this social category or exposure to the healthcare environment. Previous research has highlighted the tendency of healthcare students to identify themselves as healthcare providers from the early years of their training (Adams et al., 2006; Burford & Rosenthal-Stott, 2017; Cook & Stoecker, 2014). Nonetheless, they have not specifically investigated their stereotypical mental representation (Adams et al., 2006; Burford & Rosenthal-Stott, 2017; Cook & Stoecker, 2014). Our results indicate that healthcare students at the very beginning of their professional education may develop a stereotypical representation similar to that of the general population. Further

studies should explore these findings, delving into their replication among healthcare providers and people outside the medical field from diverse backgrounds.

Moreover, although the three subsamples differed both in sex distribution and age, the “good doctor” models estimated from each subsample were nearly indistinguishable from one another and the overall model; indicating that sample- composition differences did not affect the result in any way.

Experiment 2 provided robust validation for the effectiveness of the “good doctor” dimension by demonstrating that decisions or preferences can be influenced by stereotyping alone. The findings revealed that even in the absence of additional information about the presented identities, faces categorized as “high” on the “good doctor” dimension were more likely to be recommended than “neutral” or “low” counterparts. This result indicates that facial stimuli can be effectively manipulated along the “good doctor” dimension to trigger stereotypical associations, thereby influencing social judgments and preferences. It reinforces the idea that the “good doctor” stereotype is embodied in mental representations and can modulate decision-making processes, even without further context or information.

Moreover, these results substantiate the link between social group membership and social attributions from faces (Bagnis et al., 2019; Bin Meshar et al., 2022; Stolier et al., 2020). The findings imply that when patients first encounter a doctor, inferential processes may be triggered; influencing patients' responses based on the degree to which the doctor aligns with their stereotype of a “good doctor”. This underscores the importance of facial cues in shaping perceptions and interactions in healthcare, highlighting how closely a doctor’s appearance adheres to this stereotype can impact patient attitudes and behaviors.

The present research has some limitations worth noting. First, the three samples primarily included Caucasian individuals of Italian citizenship. Cultural context is known to shape stereotypes; therefore, this aspect may have led to a stereotypical representation reflecting the social environment our participants are exposed to, limiting the generalizability of our results. Additionally, participants were asked to judge facial stimuli according to their resemblance with a “good doctor”, without specifying any medical specialties. We are aware that physicians specialized in different medical branches may be stereotypically associated with distinct social attributes (Ashton-James et al., 2019; Harendza & Pyra, 2017; Hill et al., 2014). Nevertheless, the type of specialist participant referred to during task completion may be considered a context-dependent factor we did not intend to investigate in this study. Furthermore, the choice of keeping a generic label (e.g., a “good doctor”) is coherent with

previous similar studies on the topic (Hall et al., 2020; Necka et al., 2021). Addressing these aspects in future research will be crucial to fully capture the variability and generalizability of the 'good doctor' stereotype across cultures and medical specialties.

3.5 Conclusion

Research on first impressions in the healthcare field is becoming increasingly relevant. In the healthcare context, first impressions critically influence whether the patients will manifest a positive or negative attitude towards healthcare providers (Hall et al., 2020; Necka et al., 2021), with consequences on patients' decisions – as their preference for a doctor over another– and important medical outcomes, including pain expectations and analgesic consumption (Necka et al., 2021). It is important to note that medical outcomes, like perceived pain, may be directly impacted by factors such as facial cues (Bagnis et al., 2023; Mattarozzi et al., 2021). Moreover, in a world that is systematically pushing patients toward telemedicine and the increased use of online platforms to interact with healthcare providers (Emerick et al., 2020), the doctor's online image– such as photographs and profiles– is often the only information upon which patients can rely to make their selection or gather relevant details.

In the present study, we highlighted the social attributes – loading in the social components of ethnicity and perceived similarity, attractiveness, trustworthiness, and familiarity and professional look/intelligence– that contribute in shaping the "good doctor" stereotype and, consequently, makes a doctor more likely to be chosen over another. The novelty of our experimental manipulation lies in developing a model that not only visualizes such stereotypical representation, but offers a valid tool to investigate the impact it has on significant medical outcomes. This contributes to corroborate and expands the current knowledge on the effect of inferential processes in the medical field, a body of evidence demonstrating that such inferences not only modulate patient satisfaction but also influence the perception of pain during invasive procedures (Bagnis et al., 2023; Mattarozzi et al., 2021).

Application of this model dimension in future experimental studies holds promise for enhancing personalized medicine, improving patient satisfaction, and optimizing healthcare delivery by integrating the nuanced understanding of non-verbal communication within clinical settings.

4 From stereotypes to pain: How faces aligning with the "good doctor" mental representation modulates pain perception

4.1 Introduction

In the previous study, we identified the stereotypical mental representation of a "good doctor" and developed a model able to visualize this representation. We also explored the social attributes defining this stereotype, which clustered into four social dimensions: ethnicity and perceived similarity, attractiveness, trustworthiness, and familiarity and professional look/intelligence. While this model improved our understanding of how people visualize a "good doctor", the potential impact of such representation of relevant clinical outcomes, as pain perception, remains unexplored.

Inferences of intrinsic familiarity from facial appearance has already been identified as a possible factor modulating individuals' experience of pain (Bagnis et al., 2023; Mattarozzi et al., 2021). Building on this evidence, the present study aimed to investigate whether facial stimuli aligning with the stereotypical representation of a "good doctor" could modulate subjective pain experience, by reducing perceived threat through inferences of familiarity.

To pursue this goal, participants were exposed to facial images manipulated along the "good doctor" dimension while receiving painful electrical stimuli and were asked to rate both pain intensity and unpleasantness.

We predicted that faces aligning with the stereotypical representation of a "good doctor" – namely "high good doctor" faces– would determine placebo hypoalgesia, defined as a psycho-neuro-biological mechanism that is activated whenever there is a change in health status and capable of producing positive health changes in symptoms (e.g., pain) through mechanisms triggered by individual and social contextual factors (Bagnis et al., 2025; Colloca et al., 2019; Skuse & Gallagher, 2011). This effect was expected to be mediated by inferences of familiarity: in line with the strong correlation between the "good doctor" dimension and "familiarity" observed in Study 1, and consistent with previous evidence (Bagnis et al., 2023; Mattarozzi et al., 2021) and both the SBT (Coan & Sbarra, 2015) and the "familiar face overgeneralization" hypothesis (Zebrowitz & Montepare, 2008), highly familiar- looking faces, as "high good doctor" faces, may act as "safety signals", triggering reassurance and ultimately reducing perceived pain.

Furthermore, considering the role that genetic variations play in modulating hypoalgesic effects (Colagiuri et al., 2015; Hall et al., 2015), the present study also explored whether the pain modulatory effect of this stereotypical representation may be mediated by the genetic profile. In particular, we focused on two groups of Single Nucleotide polymorphisms (SNPs): SNPs contributing to shape pain modulation, namely i) COMT rs4680, related to the dopaminergic system and ii) FAAH rs324420, involved in the endocannabinoid system; and SNPs correlating with socio-affiliative behaviours, namely iii) TPH2 rs4570625, related to the serotoninergic system and iv) OXTR rs2254298, involved in the functioning of the oxytocinergic system.

We hypothesized that individuals carrying the COMT GG variant – typically associated with greater stress resilience (Serrano et al., 2019) and better emotion regulation (Drabant et al., 2006) – and FAAH A*variant – which have been associated with reduced anxiety (Dincheva et al., 2015; Mayo et al., 2020) – would show the strongest benefit from the presentation of faces reflecting the stereotypical representation of a "good doctor" during painful stimulation, reporting lower pain intensity and unpleasantness. Finally, we hypothesized that carriers of the T variant of TPH2– previously linked to heightened amygdala reactivity to emotional stimuli (Brown et al., 2005) – would show divergent tendencies depending on the allelic form of OXTR rs2254298. Specifically, we expect to find that individuals carrying the more prosocial C allele will exhibit stronger pain modulation in response to positive social stimuli (i.e., "high good doctor" faces), whereas carriers of the less prosocial T allele will benefit more in the absence of social salient stimuli (i.e., fixation crosses).

Notably, Heart rate (HR) and Heart Rate Variability (HRV) components were recorded throughout the experimental procedure, as physiological indexes signaling sympathetic/parasympathetic activation and variations in perceived pain (Lavigne et al., 2001; Tousignant-Laflamme et al., 2005; Treister et al., 2012). In line with previous studies testing the effect of emotional cues on physiological measures, included HR and HRV (Balconi et al., 2009; Hillmer et al., 2021)- and in coherence with the induction of positive and negative affective states by the presentation of "high" and "low good doctor" faces- we expect to find acceleration and deceleration of HR and HRV components in response to positively and negatively valenced faces, respectively.

4.2 Materials and Methods

4.2.1. Participants

205 Caucasian students enrolled in healthcare courses at the University of Bologna and nursing tutors (151 females, 54 males, age range: 18-63 years, *M*=23.50; *SD*=8.04) were recruited during lectures and invited to participate in a study on pain modulation.

65 participants were excluded from the analyses. The following exclusion criteria were adopted: i) medical/psychiatric diagnoses which may cause an altered perception of pain compared to other participants (e.g., Ehlers-danlos syndrome, obesity, anorexia, GAD) (*N*=2); ii) lack of DNA sample collection (*N*=18); iii); assumption of painkiller 12h before undergoing the task (*N*=3); iv) interruption of the task before its completion (*N*=1); v) exhibition of scarce compliance to the tasks (e.g., they did not looked at the screen during the experiment or they asked questions unrelated to the task, despite being instructed to remain silent) or lack of variability in the given responses (e.g., they consistently reported the same value on every trial for both pain intensity and unpleasantness) (*N*=41). Therefore, the final sample consisted of 140 Caucasian participants (102 females, 38 males, age range: 18-58 years, *M*=22.82; *SD*=7.70).

Sample size was established through an a-priori calculation using G*Power software (Faul et al., 2007), which indicated that assuming a small effect size (*f*=0.20) and a correlation of 0.5 among repeated measures, a minimum of 42 participants would be required to achieve a statistical power of .80 for alpha=0.05.

For the genetic analyses, COMT rs4680 (GG vs. AG vs. AA), FAAH rs324420 (CC vs. A*), TPH2 rs4570625 (GG vs. T*) and OXTR rs2254298 (CC vs. T*) were treated as fixed factors. Although validated methods for power and sample size estimation in mixed-model designs are not yet available (Guo et al., 2013), we followed the procedure described in Colloca et al., 2019 and applied power analysis for repeated-measures ANOVA as an approximation. Therefore, assuming a small effect size (f=0.15, correlation=0.50) and considering 6 different groups (GG/CC; AG/CC; AA/CC; GG/A*; AG/A*; AA/A*), a minimum of 126 participants would be required to achieve a power of 0.80 with $\alpha = 0.05$.

The study was reviewed and approved by the Bioethics Committee of the University of Bologna (Prot. No. 0388067).

4.2.2 Stimuli

8 Caucasian identities (4 females/4 males) were selected from the Chicago Face Database (Ma et al., 2015) according to their standardized appearance-based ratings of familiarity.

Specifically, Z-scores for familiarity ratings were calculated for each face based on the evaluations provided by an independent sample of 342 Italian participants (295 women, 47 men, *M*=22.73, SD=6.94 years). Faces with Z-scores closest to zero (i.e., as neutral as possible in terms of familiarity) were then selected and manipulated along the "good doctor" dimension (for further details, see Chapter 3.0) to generate the "high good doctor" (i.e., +3 SD from the mean) and the "low good doctor" (i.e., -3 SD from the mean) version of each identity. Faces that did not display potentially distracting features after the manipulation (i.e., glasses or a visible-tooth smile) were considered eligible for the experimental study.

Therefore, a total of 24 stimuli were used: 16 facial stimuli (8 identities x 2 conditions: high vs. low) and 8 fixation crosses, as a control condition. No significant differences in familiarity scores were observed among the stimuli presented during the experiment ($t(18)=.98$, $p=.33$).

Task administration occurred via E-prime software 2.0 (Schneider et al., 2002).

4.2.3 General procedure

The experiment consisted of three phases.

First, participants completed some questionnaires and underwent saliva sample collection for DNA extraction.

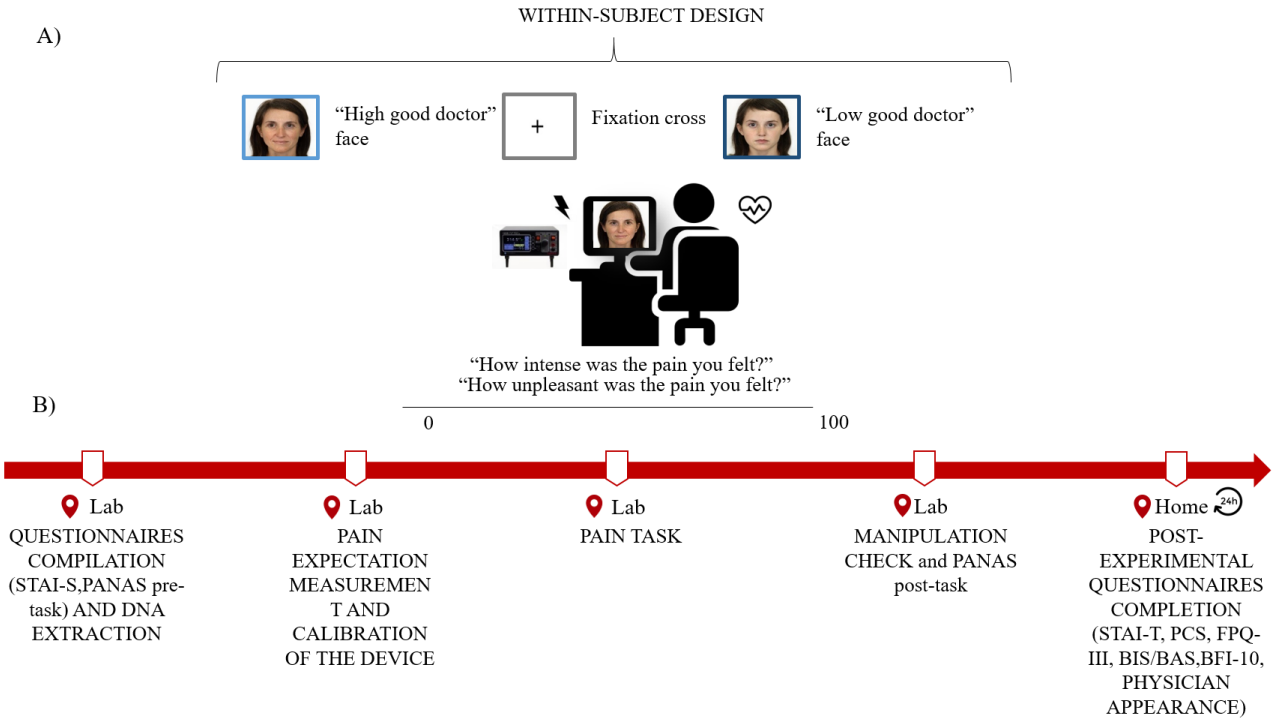
Second, they performed a pain task in which visual stimuli were paired with painful stimulations (i.e., electric impulses), while their HR and HRV components were measured. They subsequently completed a rating task of the previously displayed stimuli, as manipulation check.

Finally, participants completed additional questionnaires at home, which were sent via e-mail.

The overall experimental procedure is illustrated in Figure 4.1

Figure 4.1

Pain task design (A) and timeline of the experimental procedure (B)



Note. Panel A displays the pain task design, while panel B the timeline of the experimental procedure. The order of block conditions (high, low, fixation cross) was counterbalanced across participants.

Questionnaires compilation and DNA extraction. After providing informed consent, participants received an e-mail from the experimenter to schedule their appointment for the experimental session. On the day prior to the experimental session, participants were instructed to abstain from alcohol, nicotine, and caffeine for 24 hours and to refrain from strenuous physical activity. Upon arrival at the laboratory, participants were asked to take a seat in a room where they answered questions regarding exclusion criteria (health status, presence of medical conditions, painkiller assumption 24h prior the experimental session) and filled out two questionnaires, namely i) the State Anxiety Inventory (STAI-S, Spielberger, 1989), a measure of the state anxiety; ii) the Positive Affect and Negative Affect Scale (PANAS, Watson et al., 1988), assessing both positive and negative affective states. Once completed the questionnaires, a saliva sample was manually collected for DNA extraction and genotyping using hDNAfreeFLOQSwabs®, a non-invasive saliva swab designed for genetic applications. Participants were instructed to place the swab in their mouth for approximately 10 s, followed by gently rubbing

it against the inner cheeks and oral vestibules for about 40 s to ensure adequate sample collection. The swabs were then stored at −20 °C until analysis.

Then, two electrocardiography electrodes were applied on the right clavicle and the lowest left rib for HR recording. HR was recorded as beat-to-beat intervals in ms with the Bodyguard 2 (Firstbeat Technologies Ltd.; Parak & Korhonen, 2013).

Participants were subsequently invited to sit in front of a PC screen (screen size: 30 × 37 cm; visual angle: 2.78° × 4.28°; resolution: 1280 × 1024). Once comfortably seated upright, they were informed about the following procedure, which included i) the calibration of the electrical stimulation device (DS8R Bipolar Constant Current Stimulator Hertfordshire, UK; range 0–1000 mA), ii) the administration of pain task, and iii) the completion of a subsequent face rating task, as manipulation check. Participants were reminded that they could withdraw from the study at any time.

Pain expectation measurement and calibration of the device. To start, participants were asked to report their pain expectation (i.e., "How much pain do you expect to feel during the procedure?") using a 0 ("Not at all") -10 ("Extremely") numeric scale. Then, two electrodes connected to the DS8R were applied on the participant's non-dominant forearm and instructions about the procedures for determining tactile and pain thresholds were provided.

The tactile threshold was determined by gradually increasing the current in 0.5 mA increments until the participant reported perceiving "something" (Riello et al., 2019). Specific instructions were: "I will now begin delivering stimuli to your forearm; please indicate as soon as you feel "something"". The pain threshold, instead, was established by increasing the current in 5 mA increments from the tactile threshold until the participant first reported perceiving pain (Siedlecka et al., 2018). Specific instructions were: "Please tell me as soon as the stimulus begins to feel painful". The stimulus intensity used during the experimental task was set individually at twice the participant's pain threshold (Brączyk & Bąbel, 2021), and remained constant throughout the experiment.

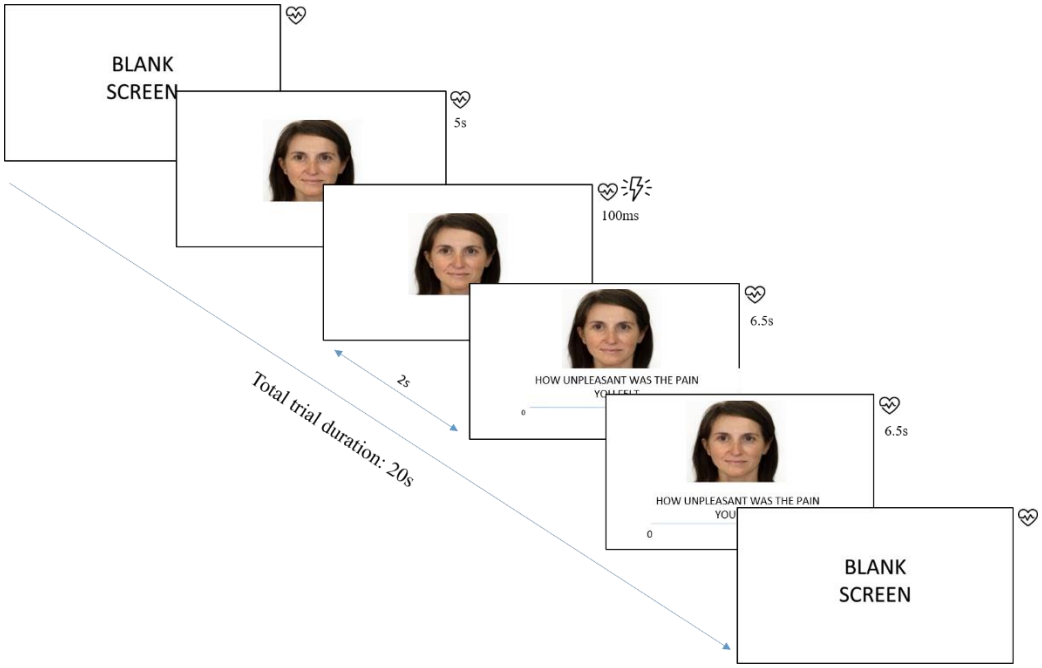
Pain Task. Following calibration, participants completed the pain task. Specifically, they were instructed to imagine themselves as about to undergo an electromyography (EMG) procedure, where the faces presented on the screen represented the doctor in charge of performing the medical procedure. Eight testing trials (6 trials without electric impulses + 2 trials with the electric impulses) were administered to allow participants to familiarize with the procedure. Then, participants received 24 electrical impulses (100 ms each, identical in intensity, set at twice the pain threshold) paired with

three types of visual stimuli: 8 “high good doctor” faces (i.e., faces manipulated at +3 SD from the mean along the “good doctor” dimension), 8 “low good doctor” faces (i.e., faces manipulated at -3 SD from the mean along the “good doctor” dimension) and 8 fixation crosses (i.e., neutral stimuli). In each trial, the visual stimulus (“high good doctor”/ “low good doctor”/ “fixation cross”) was presented for 5 seconds, followed by the delivery of an electrical impulse. Participants then had 6.5 seconds to rate both the perceived pain intensity and unpleasantness (“How intense/unpleasant was the pain youfelt?”) on a 0 (“Not at all”) to 100 (“Extremely”) numeric scale (see Figure 4.2 for trial illustration).

HR and HRV components were recorded throughout the total duration of the pain task.

Figure 4.2

Experimental trial design



Note. This is an example of “high good doctor” trial. Each trial lasted 20s. For each trial, an electric impulse was released 5s after image exposure, followed by the questions: “How intense/unpleasant was the pain you felt?”. Participants have 6.5s to respond to each question, using a 0-100 numeric scale.

The task was structured in a block design with three blocks: the “high good doctor” block, the “low good doctor” block and the “fixation cross” block. Each block consisted of 8 trials lasting 20 seconds each, for a total block duration of 160 seconds. Block order was counterbalanced across participants.

The total task duration was about 11 minutes. Once the task was terminated, the electrocardiography electrodes were removed.

Manipulation Check. After completing the pain task, participants completed the PANAS for the second time to assess their affective states, and performed an additional rating task for manipulation check. They evaluated the 8 identities previously displayed (in both their “high good doctor” and “low good doctor” versions) as well as 8 new identities (in both their “high good doctor” and “low good doctor” versions) not presented during the task, for a total of 32 facial stimuli. For each face, participants rated how much it resembled a “good doctor” on a 0 (“Not at all”) to 100 (“Extremely”) numeric scale.

Two practice trials were administered to familiarize with the procedure. Then, facial stimuli were presented one after the other, together with the question “To what extent does this face look like a “good doctor” to you?” in a randomized order.

Post-experimental questionnaire completion. Finally, participants were instructed to complete additional questionnaires to measure personality traits. Questionnaires were sent via e-mail and should be completed within 24 hours of the experimental session. The assessed personality traits included: i) trait anxiety (STAI-T; Spielberger, 1989, Italian adaptation by Pedrabissi & Santinello, 1989; possible range: 20-80, clinical cut off: 40), pain catastrophizing (Pain Catastrophizing Scale (PCS); Sullivan et al., 1995, Italian adaptation by Monticone et al., 2012; possible range: 0- 52), fear of pain (Fear of Pain Questionnaire; FPQ-III; McNeil & Rainwater, 1998, Italian adaptation by Tella et al., 2019; possible range: 30-150), the Big Five traits– extraversion, openness, agreeableness, conscientiousness, and emotional stability (BFI-10;Rammstedt & John, 2007; possible range: 2- 10 for each scale); and the behavioral inhibition/activation (BIS/BAS Scale; Carver & White, 1994, Italian adaptation by Leone et al., 2001; possible range BIS: 7- 28; Drive scale: 4-16; Fun: 4-16; Reward: 5-20). The influence of the physician’s appearance on the participant was also assessed using the questionnaire described in Aitken et al., 2014. Importantly, we measured both the importance of the physician’s physical appearance in terms of trust (item: “How important is the physician’s appearance to you in terms of trust?” on a scale from 1 (“Not at all”) to 5 (“Extremely”) and the importance of physician’s grooming (item: Please rank from 1 (more important) to 4 (least important)

the following physician's characteristics: personal grooming, kindness, empathy, clinical knowledge. Participants who ranked personal grooming as the most important characteristic were coded as 1; all others were coded as 0).

All participants were subsequently fully debriefed regarding the study's aims, manipulated variables, hypotheses, and expected outcomes. If the participant was a student, they were awarded credits for their participation.

4.2.4 DNA genotyping

Saliva samples were stored at -20°C until DNA extraction, which occurred within 10 days from collection to avoid degradation. Samples were then processed using the "Isolation of total DNA from small volumes of saliva" protocol, according to the manufacturer's instructions (DNeasy Blood & Tissue Kit QIAGEN); and the COMT rs4680, FAAH rs324420, TPH2 rs4570625 and OXTR rs2254298 polymorphisms were genotyped through a standard Polymerase Chain Reaction (PCR). Qubit™ dsDNA hs Assay Kits (100 assays; cod: Q32851, thermo scientific) and NanoDrop™ 2000 (thermo scientific) were used for DNA quantification.

4.2.5 Data analyses

To assess the effectiveness of faces manipulation along the "good doctor" dimension, a repeated-measure ANOVA was performed.

Pain measures: pain intensity and unpleasantness. Correlational analyses were conducted between mean pain intensity and unpleasantness in each condition and participants' pain expectations.

Linear mixed-effects models (LMMs) were conducted separately for the dependent variables pain intensity and unpleasantness ratings. For the primary analyses, Condition ("high good doctor"/ "low good doctor"/ "fixation cross") was entered as a fixed effect. Models included a random intercept for Participant. Fixed effects were evaluated via type-III ANOVAs using maximum likelihood. Exploratory analyses tested Condition × Participant Sex and the three-way interaction Condition × Participant Sex × Stimulus Sex.

Results are detailed in the Supplementary Materials (see subchapter 8.1.1).

To further examine whether the effect of Condition interacted with participants' genetic profile, we fitted factorial LMMs including Condition x Genotype interactions for each of the genetic

polymorphisms of interest, namely COMT rs4680 (AA vs. AG vs. GG), FAAH rs324420 (CC vs. A*), TPH2 rs4570625 (GG vs. T*) and OXTR rs2254298 (CC vs. T*). Models included Condition × Gene x Gene interactions.

Analyses focused on two combinations of genes per model: i) COMT rs6480 and FAAH rs324420, which have previously been investigated in placebo hypoalgesia (Colloca, et al., 2019); and ii) TPH2 rs4570625 and OXTR rs2254298, which have never been examined in combination within the context of placebo hypoalgesia, but have individually been linked to anxiety (Laas et al., 2017; Thompson et al., 2011) and socio-affective processes such as emotion regulation (Inoue et al., 2010; Waider et al., 2011) and social cognition (Brüne, 2012).

Effect sizes were quantified with marginal and conditional R^2. Significant effects were further explored with estimated marginal means and Bonferroni-corrected pairwise contrasts.

Physiological measures: mean HR and HRV components. To test the effect of Condition on the mean HR and the parasympathetic and sympathetic components of HRV (i.e., RMSSD, VLF, LF, HF, HF/LF) calculated for each block, we applied the following exclusion criteria on our sample (*N*=140): i) missing HR recording during the task performance (*N*=18); ii) Z scores for mean HR and HRV components <-3 and >3 (*N*=14). Therefore, analyses on physiological measures were conducted on a sample of 108 participants (78 females, 30 males, age range: 18-58; *M*=21.52; *SD*=5.14).

Heart rate was recorded as beat-to-beat intervals in ms with the Bodyguard 2 (Firstbeat Technologies Ltd.;Parak & Korhonen, 2013). HRV was derived by computing the root mean square of successive beat-to-beat interval differences (rMSSD), which reflects vagal regulation of HR with reduced susceptibility to respiratory influences (Laborde et al., 2017; Penttilä et al., 2001). Analyses were performed using Kubios HRV software (Tarvainen et al., 2014) which uses an advanced detrending method based on the formulation of smoothing priorities (Tarvainen et al., 2002). Artifacts and ectopic beats were corrected using Kubios HRV software by two different methods: (1) A threshold-based correction, which is based on the comparison of each RR interval value to a local average interval; (2) an automatic correction, where artifacts are detected from a time series consisting of the differences between successive RR intervals.

Analyses on the effect of Condition on the mean HR and HRV components for each block were performed through repeated-measure ANOVA. Effect sizes were quantified with partial eta squared (η^2p).

Questionnaires. A paired samples t-test was performed on PANAS scores collected before and after the pain task to examine whether differences in reported positive and negative affective states emerged. Moreover, descriptive analyses were conducted on all questionnaires scoring to provide an overview of the sample characteristics. Finally, correlational analyses with the main dependent variables were also performed, to explore possible associations between personality traits and the experimental variables. Questionnaires scorings were not used as covariates in the main analyses due to the small sample size.

From the initial sample (N=140), 12 participants were excluded from analyses due to the lack of questionnaires completion (N=8) or Z scores <-3 and >3 (N=5). Therefore, these exploratory analyses were conducted on a sample of 127 participants (92 women, 35 men, range age: 18-58 years, M=23.09, SD=8.03).

All analyses were performed in SPSS (version 25.0; Chicago, IL) and R (version 4.4.1; 2024-06-14 ucrt; Posit, PBC, Boston, MA, USA), using the lme4, emmeans and performance packages.

The alpha level for all analyses was set to p<.05.

4.3 Results

The analysis confirmed the manipulation was effective (F(1,139)=486.21, p<.001, η^2=.78), revealing that participants judged "high good doctor" faces (M=70.79, SD=11.41) as significantly more similar to the stereotypical representation of a "good doctor" than "low good doctor" faces (M=45.42, SD=13.85).

Results on pain task, instead, are presented separately for pain and physiological measures.

4.3.1 Pain measures: pain intensity and unpleasantness

4.3.1.1 Main effect of Condition

Correlational results evidenced significant positive correlations between participant's pain expectation and pain intensity across all conditions (r_{high}=.25, p<.01; r_{low}=.29, p<.001; r_{cross}=.21, p<.05), as well as between pain expectations and unpleasantness in the "low good doctor" condition (r=.21, p<.05)

Linear mixed-effects models (LMMs) revealed a significant main effect of Condition ("high good doctor" / "low good doctor" / "fixation cross") on both pain intensity and pain unpleasantness ratings (pain intensity: $F(2,3218)=7.01$, $p<.001$, $R^2c=.75$, $R^2_m=.001$; pain unpleasantness: $F(2,3218)=8.96$, $p<.001$, $R^2c=.72$, $R^2_m=.002$).

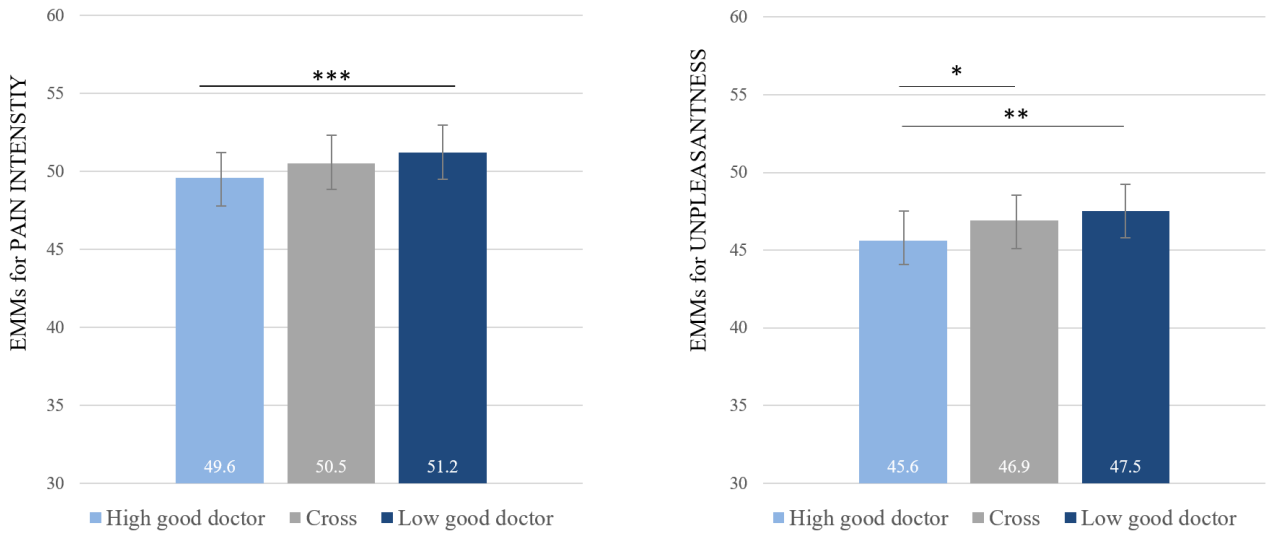
Bonferroni- adjusted pairwise comparisons indicated that participants reported lower pain intensity in the "high good doctor" (*EMM*=49.60, 95% *CI*[46.80, 52.50]) compared to the "low good doctor" condition (*EMM*=51.20, 95% *CI*[48.30, 54.00], $p<.001$).

Similarly, they reported lower unpleasantness in the "high good doctor" condition (*EMM*=45.6, 95% *CI*[42.7, 48.6]) compared to both the "low good doctor" (*EMM*=47.50, 95% *CI*[44.60, 50.50], $p<.001$) and the fixation cross (*EMM*=46.90, 95% *CI*[44.00, 49.90], $p<.05$) conditions.

Results are displayed in Figure 4.3.

Figure 4.3

Estimated marginal means (EMMs) of pain intensity and unpleasantness across the "high good doctor", "low good doctor" and fixation cross conditions.



Note. Error bars represent the standard error of the EMMs**.** * p<.05; **$p<.01$; ***$p<.001$

No other significant differences emerged. The complete output for the linear mixed-effect models and all the results are displayed in Supplementary materials (see subchapter 8.1.2, Tables S8.5 and S8.6).

4.3.1.2 Condition x Gene x Gene interactions

Condition × COMT rs4680 (AA vs. AG vs. GG) × FAAH rs324420 (A* vs. CC). Participants were distributed across the genotype groups as follows: AA/A* (*N*=9), AA/CC (*N* =15), AG/A* (*N* =33), AG/CC (*N* =40), GG/A* (*N* =10), and GG/ CC (*N* =33).

The three-way interaction between Condition, COMT rs4680 and FAAH rs324420 was significant for both pain intensity and unpleasantness (pain intensity: $F(4,3220)=7.45$, $p<.001$, $R^2c=.75$, $R^2_m=.03$; pain unpleasantness: $F(4,3220)=11.87$, $p<.001$, $R^2c=.72$, $R^2_m=.04$), indicating that the effect of Condition varied depending on participants' genetic profile.

Bonferroni-adjusted pairwise comparisons showed that GG/A* carriers reported significant lower pain intensity in the "high good doctor" (*M*=45.50, 95% *CI*[34.70, 56.20]) compared to the "low good doctor" (*M*=49.40, 95% *CI*[38.60, 60.10], $p<.05$) and fixation cross conditions (*M*=54.10, 95% *CI*[43.40, 64.80], $p<.001$). A significant difference also emerged between the "low good doctor" (*M*=49.40, 95% *CI*[38.60, 60.10], $p<.05$) and the fixation cross condition (*M*=54.10, 95% *CI*[43.40, 64.80], $p<.001$) ($p<.01$).

Similarly, AA/CC carriers reported significant lower pain intensity in the "high good doctor" (*M*=43.00, 95% *CI*[34.20, 51.80]) compared to the "low good doctor" (*M*=47.60, 95% *CI*[38.90, 56.40], $p<.001$) and fixation cross conditions (*M*=47.40, 95% *CI*[38.60, 56.10], $p<.01$).

Results on pain unpleasantness mirrored these findings.

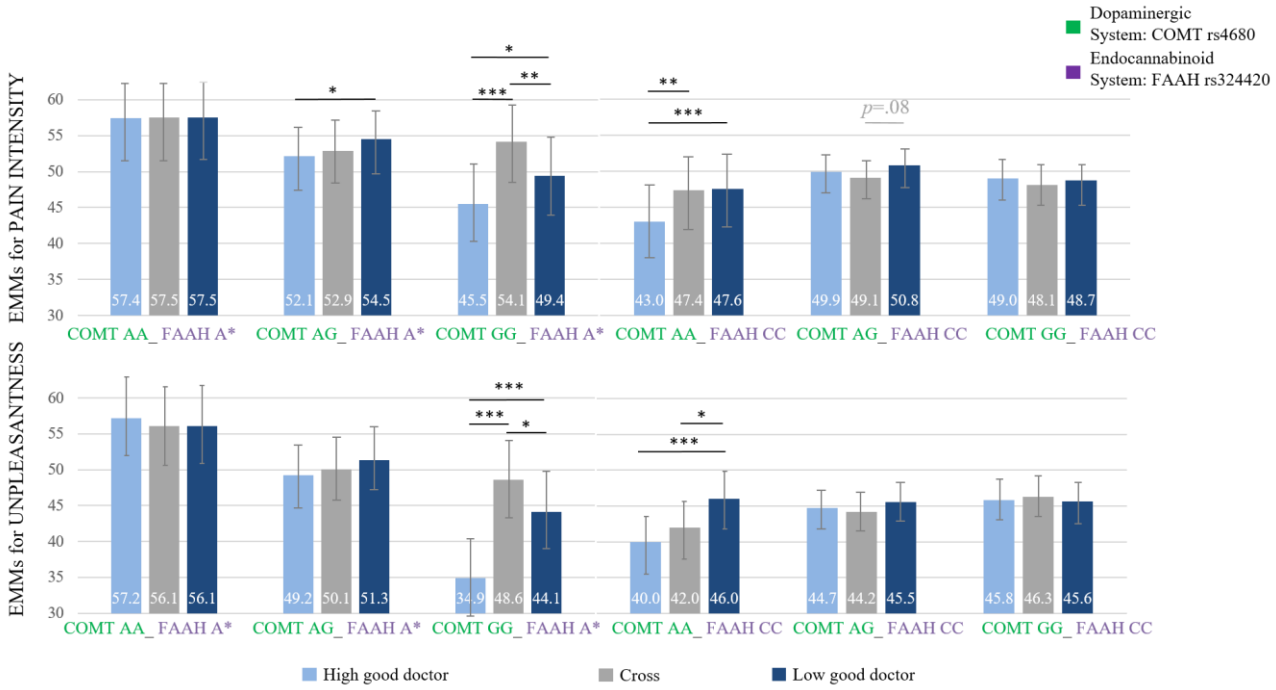
GG/A* carriers reported significant lower unpleasantness in the "high good doctor" (*M*=34.90, 95% *CI*[24.00, 45.90]) compared to the "low good doctor" (*M*=44.10, 95% *CI*[33.20, 55.10], $p<.001$) and the fixation cross conditions (*M*=48.60, 95% *CI*[37.60, 59.60], $p<.001$). A significant difference also emerged between the "low good doctor" and the fixation cross condition ($p<.05$), with the latter being associated with the higher unpleasantness ratings.

Among AA/CC carriers, unpleasantness was significantly lower in the high good doctor (*M*=40.00, 95% *CI* [31.00, 48.90]) compared to the low good doctor condition (*M*=46.00, 95% *CI*[37.10, 55.00], $p<.001$). A significant difference also emerged between the "low good doctor" and the fixation cross condition (*M*=42.00, 95% *CI*[33.10, 51.00], $p<.05$).

Results are displayed in Figure 4.4.

Figure 4.4

EMMs of pain intensity and unpleasantness in the "high good doctor", "low good doctor" and fixation cross conditions in participants with different variants of COMT rs4680 and FAAH rs324420.



Note. Error bars represent the standard error of the EMMs. *p<.05; **p<.01; ***p<.001

No other significant differences emerged across genotypes or conditions. The complete output for the linear mixed-effect models and all results are displayed in Supplementary materials (see subchapter 8.1.3, Table S8.7 and S8.8).

Condition × TPH2 rs4570625 (GG vs. T*) × OXTR rs2254298 (CC vs. T*). Participants were distributed across the genotype groups as follows: GG/CC(N=54), GG/T* (N=29), T/CC (N=45), T*/T* (N=12).

As illustrated in Figure 4.5, the three-way interaction between Condition, TPH2 rs4570625 and OXTR rs2254298 was significant for both pain intensity and unpleasantness (pain intensity: F(2, 3220)=25.41, p<.001, R^2c =.75, R^2_m=.02; pain unpleasantness: F(2, 3220)=19.88, p<.001, R^2c =.72, R^2_m=.02).

Bonferroni- adjusted pairwise comparisons revealed that carriers of the GG/CC genotype reported marginally significant lower pain intensity in the "high good doctor" condition (*M*=51.50, 95% *CI*[46.90, 56.10]) compared to the "low good doctor" condition (*M*=53.10, 95% *CI*[48.5, 57.7],*p*=.06). T*/CC carriers also showed a marginally lower pain intensity in the "high" (*M*=47.0, 95% *CI*[42.0, 52.1]) versus the "low good doctor" condition (*M*=48.6, 95% *CI*[43.6, 53.7], *p*=.09), and significantly lower ratings in the "high good doctor" condition compared to the fixation cross (*M*=49.5, 95% *CI*[44.4, 54.6], *p*<.01). Conversely, carriers of the T*/T* genotype reported significantly lower pain intensity when viewing fixation crosses (*M*=47.9, 95% *CI*[38.1, 57.7]) rather than both "high" (*M*=56.4, 95% *CI*[46.6, 66.2], *p*<.001) and "low good doctor" faces (*M*=58.2, 95% *CI*[48.4, 68.0], *p*<.001).

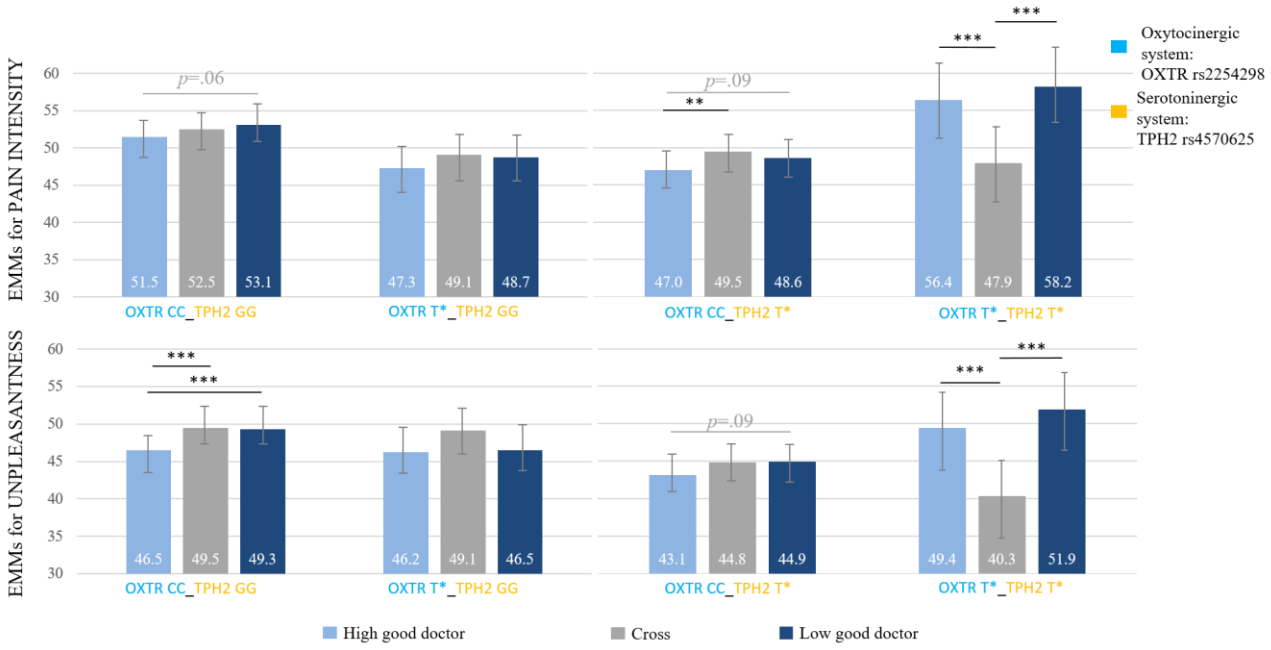
Results on pain unpleasantness partially replicate those for pain intensity.

Specifically, GG/CC reported lower unpleasantness in the "high good doctor" (*M*=46.50, 95% *CI*[41.80, 51.30]), compared to both the "low good doctor" (*M*=49.30, 95% *CI*[44.50, 54.00], *p*<.001) and the fixation cross conditions (*M*=49.50, 95% *CI*[44.80, 54.30], *p*<.001). In contrast, T*/T* reported significantly lower unpleasantness ratings when fixation crosses (*M*=40.30, 95% *CI*[30.20, 50.40]), rather than "high" (*M*=49.40, 95% *CI*[39.30, 59.50], *p*<.001) and "low good doctor" faces (*M*=51.90, 95% *CI*[41.80, 62.00], *p*<.001), were presented.

Results are displayed in Figure 4.5.

Figure 4.5

EMMs of pain intensity and unpleasantness in the "high good doctor", "low good doctor" and fixation cross conditions in participants with different variants of OXTR rs22543298 andTPH2 rs4570625.



Note. Error bars represent the standard error of the EMMs. *p<.05; **p<.01; ***p<.001

No other significant differences were observed. The complete output for the linear mixed-effect models and all results are displayed in Supplementary materials (see subchapter 8.1.4, Table S8.9 and S8.10).

4.3.2 Psychophysiological measure: Mean HR and HRV components during pain

4.3.2.1 Main effect of Condition

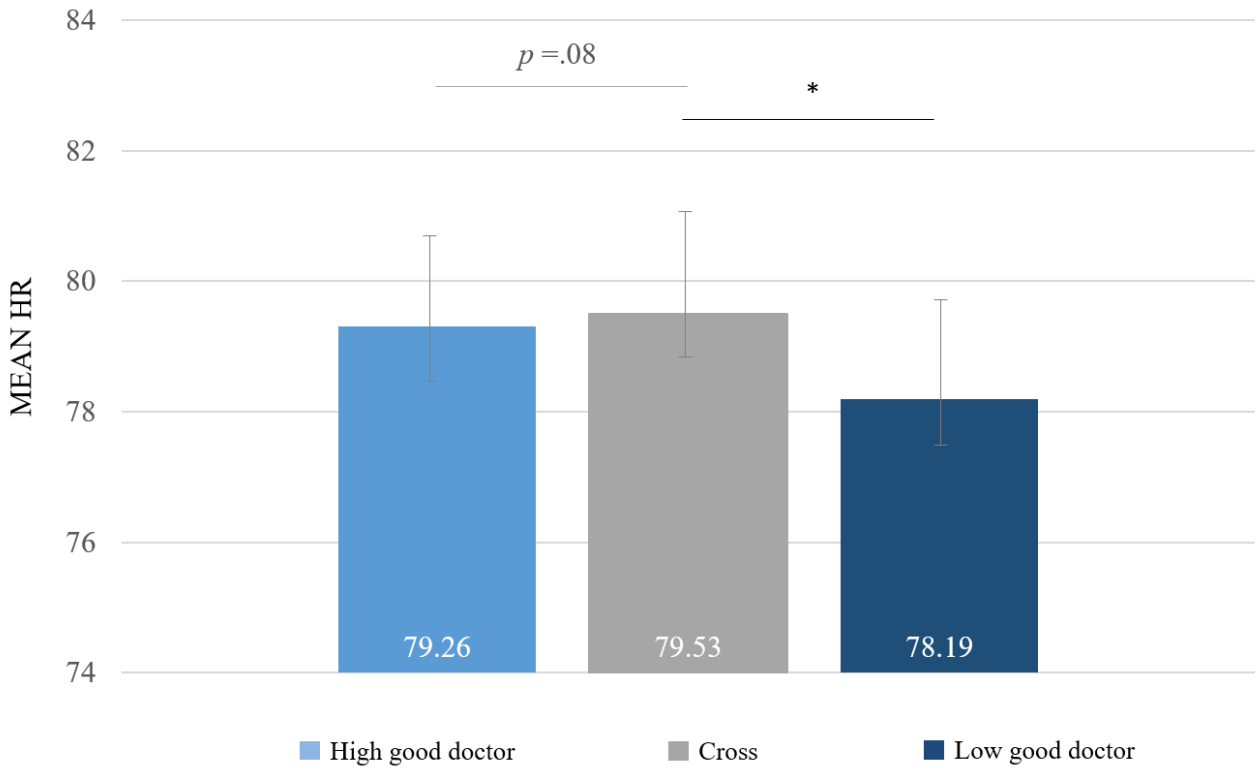
A repeated-measure ANOVA evidenced a main effect of Condition on Mean HR for each block (F(1.76,188.76)=4.96, p<.01, η^2_p=.05).

Specifically, we observed significantly lower mean HR in the "low good doctor" (M=78.19, SD=11.46) compared to fixation cross condition (M=79.53, SD=10.55; p<.05). A tendency toward a significant difference was also observed between mean HR in the "low good doctor" and the "high good doctor" condition (M=79.26, SD=10.39, p=.08), which was associated with higher HR.

Results are displayed in Figure 4.6.

Figure 4.6

Mean HR in the "high good doctor", "low good doctor" and "fixation cross" conditions.



Note. Error bars represent the standard error of the EMMs. *p<.05

No other significant results were observed for HRV components (i.e., RMSSD, VLF, LF, HF, HF/LF).

Analyses on the main effect of Condition on HRV components and the corresponding results are reported in the Supplementary materials (see subchapter 8.1.5, Table S8.10 and S8.11).

4.3.3 Questionnaires

A paired samples t-test on PANAS scoring collected before and after undergoing pain task showed that although no differences in positive affective states were observed (t(126)=0.53, p=.60), participants reported significantly higher negative affect before (M=16.51;SD=4.97) than after (M=13.29, SD=3.92) the pain task.

Descriptive analyses on questionnaires scorings also evidenced that participants generally reported higher positive vs. negative affective states.

Results of descriptive analyses are displayed in Table 4.1.

Table 4.1

Descriptive statistics for the questionnaires administered during the whole experimental procedure.

	Minimum	**Maximum**	***M***	***SD***
PANAS				
Positive Affect pre pain	23	44	32.54	4.34
Negative Affect pre pain	10	32	16.51	4.97
Positive Affect post pain	20	43	32.38	4.81
Negative Affect post pain	10	28	13.29	3.92
STAI				
Trait Anxiety	23	58	45.91	9.97
State Anxiety	28	72	37.17	7.22
FPQ-III				
Fear of Pain	34	116	75.94	15.25
PCS				
Pain Catastrophizing	1	46	22.34	8.79
BFI				
Agreeableness	2	10	6.08	1.67
Conscientiousness	4	10	7.32	1.38
Emotional Stability	2	10	5.87	1.89
Extraversion	2	10	5.86	1.75
Openness	2	10	6.74	1.84
BIS/BAS				
Behavioural Inhibition	14	28	21.88	3.39
Drive	7	16	11.76	1.99
Fun	7	16	11.78	1.96
Reward	12	20	17.58	1.82
Importance trust	1	5	3.30	1.16

Correlational analyses evidenced that mean pain intensity in the "high good doctor" and "low good doctor" conditions correlated with post-pain Negative Affect (r_{high}=.18, p<.05; r_{low}=.22, p<.05), suggesting that the higher the experienced pain, the stronger the negative affect elicited. Pain intensity across all conditions showed positive correlations with the Fear of Pain mean scorings (r_{high}=.18, p<.05; r_{low}=.27, p<.01; r_{cross}=.22, p<.01), evidencing a general tendency to experience more pain in individuals with higher fear of pain. Finally, pain intensity in the "low good doctor" condition positively correlated with the importance participants associate to the physician's appearance (r_{low}=.20, p<.05), so that participants who attributed greater importance to the physician's physical appearance in terms of trust experienced higher pain ratings when exposed to "low good doctor" faces or in the absence of facial cues.

Results on pain unpleasantness partially replicated those observed for pain intensity.

Analyses revealed a positive correlation between mean unpleasantness in the "low good doctor" condition and post-pain Negative Affect (r=.19, p<.05), evidencing more intense negative affects when higher unpleasantness was experienced in this condition. Moreover, positive correlation emerged between the unpleasantness in the fixation cross condition and the Behavioural Activation Scale of Reward (r=.19, p<.05), suggesting that the stronger the participant's sensitivity to rewards, the higher the unpleasantness experienced when no facial cues were administered. Lastly, positive correlations emerged between the mean unpleasantness across all conditions and the importance attributed to the physician's appearance in terms of trust (r_{high}=.18, p<.05; r_{low}=.23, p<.01; r_{cross}=.24, p<.01)– indicating greater unpleasantness when more importance was attributed to the physician's physical appearance, regardless of whether the facial cue was presented or whether it resembled a "good doctor".

Correlations between pain measures and questionnaires scorings are displayed in Table 4.2.

Table 4.2

Correlations between pain measures (i.e., pain intensity and unpleasantness) across all conditions ("high good doctor"/ "low good doctor" / fixation cross) and questionnaires

	Pain intensity High good doctor	Pain intensity Low good doctor	Pain intensity Cross	Unpleasantness High good doctor	Unpleasantness Low good doctor	Unpleasantness Cross
PANAS						
Positive Affect pre pain	.06	.16	.10	.01	.15	.06
Negative Affect pre pain	.14	.17	.14	.11	.14	.12
Positive Affect post pain	.02	.06	.04	-.11	.01	-.03
Negative Affect post pain	.18*	.22**	.14	.16	.19*	.10
STAI						
Trait Anxiety	.08	.08	.03	.14	.16	.09
State Anxiety	-.06	-.05	-.02	-.01	.03	.02
FPQ						
Fear of Pain	.18*	.27***	.22*	.05	.12	.10
PCS						
Pain Catastrophizing	.05	.02	.04	-.02	-.02	-.05
BFI						
Agreeableness	-.01	-.04	.04	-.01	-.001	.06
Conscientiousness	.06	.04	.01	.11	.05	.07
Emotional Stability	-.10	-.07	-.09	-.07	-.05	-.05
Extraversion	.05	.06	.05	.09	.11	.09
Openness	.03	.01	.04	.01	-.001	.02

BIS/BAS						
Behavioural Inhibition	.08	.06	.07	.10	.12	.03
Drive	.02	.09	.08	-.11	-.09	-.04
Fun	-.06	-.05	-.003	-.02	-.01	.06
Reward	.06	.06	.14	.09	.12	.19*
Importance trust	.12	.20*	.17	.18*	.23**	.24**

Note. *p<.05, **p<.01, ***p<.001.

Correlations between questionnaires scoring are displayed in the Supplementary Materials (see subchapter 8.1.6, Table S8.12)

4.4 Discussion

First impressions play a critical role in healthcare, with previous research (e.g., Bagnis et al., 2023; Mattarozzi et al., 2021; Necka et al., 2021) highlighting how social attributions derived from faces– particularly the perception of familiarity– can influence pain modulation. Yet, no previous study has directly investigated how inferences of social traits people associate with the stereotypical representation of a "good doctor"- as familiarity and professional look/intelligence, but also ethnicity and perceived similarity, attractivevness and trustworthiness- modulate pain.

The present study investigated the extent to which faces aligned with the stereotypical representation of a "good doctor" can modulate the subjective experience of pain. Furthermore, it examined whether individual differences in the genetic profile may influence this modulation.

Consistent with our prediction, we found that faces manipulated to embody the physical features and social attributions associated with the "good doctor" stereotype- familiarity and professional look/intelligene among others- reduced perceived pain intensity and unpleasantness when presented during a pain induction procedure. The pattern of results reveals that the social attributes linked to the "good doctor" stereotype exert a more significant effect on the affective component of pain (Price, 2000), particularly unpleasantness, compared to pain intensity, considered instead as the cognitive component of pain. The effect on perceived pain intensity was indeed modest, appearing only in comparisons between high and low stereotypical faces and not differing significantly from the control condition. In contrast, the influence on unpleasantness was more robust and consistent with our hypothesis.

This pattern aligns with the notion that perceived unpleasantness, more than intensity, is strongly influenced by the social context in which the painful stimulation occurs (Fields, 1999; Leknes & Bastian, 2014). Depending on the situational context and on what the perceived pain signals, it may acquire different meanings or connotations, prompting individuals either to endure the situation or to withdraw from it (Leknes & Bastian, 2014). It follows that faces associated with the "good doctor" stereotype likely activate social and emotional processing pathways that reduce the perceived threat linked to the painful experience, thereby influencing the overall experience of pain– and more strongly, its motivational- affective dimension of unpleasantness (Fields, 1999). This reduction emphasizes the active role of contextual elements in shaping sensory perception coupled with the inherent emotional components of the pain experience, contributing to a pronounced reduction in unpleasantness.

This finding aligns with a broader body of research on the social modulation of pain (Krahé et al., 2013), emphasizing the influence of the presence of a conspecific– particularly when socially or affectively relevant– on pain perception (Eisenberger et al., 2011). More specifically, both laboratory studies (Bagnis et al., 2023) and research conducted in clinical settings (Mattarozzi et al., 2021) have demonstrated that patients' inferences of familiarity based on healthcare providers' facial appearance can attenuate pain perception. As stated by the "familiar face overgeneralization" hypothesis (Zebrowitz & Montepare, 2008) and consistent with SBT (Coan & Sbarra, 2015), high familiar-looking faces may trigger positive emotions (Dotsch et al., 2016; Zebrowitz et al., 2007), signalling contextual safety. Building on this perspective, in a threatening situation as the exposure to painful stimulation, faces resembling the stereotypical representation of a "good doctor" – which appear familiar– may serve as "safety signal", inducing a feeling of "security" that reduce anxiety and, consequently, pain (Coan et al., 2006; Krahé et al., 2013; Vlaeyen et al., 2009). Additionally, perceived competence and warmth have emerged as reducing patients' expectations of post-procedural pain and likelihood of analgesic consumption (Necka et al., 2021), also affecting their preference for one doctor over another. Therefore, since the effect of facial-based inferences are context specific (Hehman et al., 2017), it is possible that faces which appear not only familiar, but also competent and professional – as "high good doctor" faces– may be particularly effective in signaling safety in a clinical context, as they embody both familiarity and the social traits of the professional role most relevant – and most useful – to the situation at hand.

It is important to consider our findings regarding heart rate (HR), which provide additional insight into the immediate physiological responses elicited by our stimuli. Specifically, we found that

participants exhibited lower HR when faces conveying negatively social attributions, such as “low good doctor” faces, were displayed. These results are consistent with previous findings collected in both children (Hillmer et al., 2024) and adults (Balconi et al., 2009) exposed to negative visual stimuli, highlighting the induction of negative affect. Indeed, the exposure to negative stimuli may prompt greater initial orienting and attentional resources needed for context evaluation (Pastor et al., 2008), which may cause HR deceleration. In the context of our study, the lower HR observed in response to "low good doctor" faces may reflect an initial vigilance response and increased allocation of attentional resources to assess potential threat or discomfort. This physiological response underscores the salience of social cues in shaping both cognitive and affective processes, with the initial negative evaluation of these faces potentially influencing subsequent pain perception. In line with this perspective, our findings are consistent with research emphasizing the role of contextual factors in modulating pain perception via placebo mechanisms (Bagnis et al. 2025;Colloca & Barsky, 2020). The faces associated with the "good doctor" stereotype can be seen as salient social stimuli that activate these mechanisms, influencing pain perception through expectations, conditioning, and the release of endogenous opioids and other neurobiological pathways. These contextual factors can shape an individual's subjective experience of pain, underscoring the potential for social cues to influence pain perception in healthcare settings (Atlas, 2021).

Our findings revealed a complex interplay between genetic factors and the social context provided by the "good doctor" stereotype, highlighting individual differences in the modulation of pain. Specifically, the impact of our experimental condition on both pain intensity and unpleasantness varied significantly depending on participants' genetic profiles, as determined by COMT rs4680, related to dopaminergic system, and FAAH rs324420, related to endocannabinoid system. This suggests that the degree to which individuals experienced reduced pain and unpleasantness in response to the "good doctor" faces was contingent upon their unique combination of these genetic variations, indicating that genetic predispositions related to dopaminergic and endocannabinoid systems play a modulatory role in the social context of pain.

Specifically, we found that carriers homozygote for the G allele of SNP COMT rs4680, and at least one A allele for the SNP FAAH rs324420, showed the strongest socially induced pain modulation, reporting lower pain intensity and unpleasantness when the stereotypical representation of a “good doctor” was presented together with electric impulses. These results may be partially rooted in the COMT warrior/worrier model (Goldman et al., 2006; Serrano et al., 2019), positing that AA and GG alleles lead to different reactions in threatening contexts due to differences in tonic dopamine levels.

While the GG carriers– the "warriors" – manifest stronger stress resilience, probably due to lower baseline dopaminergic levels, AA carriers– the "worriers" – tend to be adversely affected by higher baseline dopamine levels, which will lead to heightened reactivity to aversive stimuli (Kilford et al., 2015). Therefore, it is likely that in stressful situations– as painful situations– the presentation of social salient stimuli as faces resembling a "good doctor" may act as particularly effective "safety signals" for these individuals, thereby buffering pain. This effect may be further enhanced by the combined presence of A allele of SNP rs324420, which has been related to decreased threat- related amygdala reactivity and consequent reduced anxiety levels (Hariri et al., 2009), as well as better emotion regulation (Gärtner et al., 2019) and increased reward sensitivity (Hariri et al., 2009); leading to the strongest modulatory effect.

Albeit more modest, a modulatory effect was also observed in carriers of COMT AA and FAAH CC genotype. This genetic profile has been already evidenced as more sensitive to placebo hypoalgesia compared to other genetic combinations (Colloca et al., 2019): COMT AA carriers have often emerged as exhibiting higher pain sensitivity (Schmahl et al., 2012; Vargas-Alarcón et al., 2007) and stronger placebo responsiveness compared to G carriers (Hall et al., 2012; Yu et al., 2014); and these effects appeared to be reinforced by the co-presence of C alleles of FAAH rs324420, which have been evidenced as sensitive to placebo hypoalgesia as well (Peciña et al., 2014). Nonetheless, due to the social valence of the stimuli presented in the study, it is likely that individuals with a more emotionally-regulated, reward sensitive combination of SNPs- as the COMT GG and FAAH AA combination– may have responded more adaptively, showing a stronger placebo hypoalgesia compared to others.

Research on pain modulation has also emphasized the potential role of genetic polymorphisms in systems regulating socio-affiliative behaviours – as the serotoninergic and oxytocinergic systems (Colagiuri et al., 2015; Kessner et al., 2013). Our results corroborate this hypothesis, showing that variations in polymorphisms involved in these systems influence both the intensity and the direction of the modulatory effects elicited by socially salient stimuli on pain perception.

Specifically, we found that individuals homozygus for the G variant of the TPH2 rs4570625– associated to normal serotonin synthesis (Chen et al., 2008; Scheuch et al., 2007) – and the C variant of OXTR rs2258498– an intronic SNP in the gene encoding the oxytocin receptor– benefited the most from the presentation of faces resembling the "good doctor" stereotype; reporting the lowest ratings of both pain intensity and unpleasantness compared to others genotype groups. Conversely, carriers of the T variants of the same SNPs considered exhibited an opposite tendency, with reduced pain

intensity and unpleasantness in the absence of facial stimuli. These results may be explained by the functional roles of the SNPs involved. T carriers of the SNP TPH2 rs4570625 seem to be more environmentally sensitive than GG carriers, evidencing a stronger amygdala response to both positively and negatively valenced faces (Brown et al., 2005). Moreover, the T variant of the SNP OXTR rs2254298 has been linked to increased risk of autism (LoParo & Waldman, 2015), depression and anxiety disorders (Brüne, 2012; Thompson et al., 2011), as well as stronger association with social impairment during early age (Parker et al., 2014).

It derives that individuals carrying the most prosocial variants of these genes may derive greater benefits when painful stimulations are administered in a social context, as the presence of familiar-looking conspecifics. Contrarily, carriers of variants associated with socio- affective impairments and lower prosociality may be more negatively impacted byothers' presence, experiencing heightened anxiety and, in turn, pain.

Moreover, in an experimental placebo analgesia model, intranasal oxytocin was found to enhance the placebo analgesic response (Kessner et al., 2013), which is also enhanced by patient-physician communication. This suggests that oxytocinergic and serotoninergic systems are critical to socially-induced placebo hypoalgesia, resulting in clinical benefits.

This study has some limitations that should be acknowledged. While the experimental design allowed for the detection of significant effects across the entire sample, the limitations primarily pertain to the results concerning genetic modulation. Specifically, the sample size for the genetic analyses is relatively small, and although significant effects emerged, comparisons between groups with fewer than ten participants should be interpreted with caution and require replication. Nonetheless, we argue that these findings may not only reinforce previous evidence, but also extend current knowledge on how genetic variability may influence placebo responsiveness to social stimuli, as faces, while acknowledging the need for larger, more diverse studies.

Additionally, the sample consisted predominantly of Caucasian women, which raises concerns about the generalizability of all results to the broader population. Nevertheless, as reported in the Supplementary Materials (see subchapter 8.1.1), no differences were observed between female and male participants for either pain intensity or unpleasantness ratings, suggesting that the reported effects are not explained by sex-related variability. Furthermore, the "good doctor" dimension was developed and validated mainly on the basis of ratings provided by Caucasian participants, and subsequently validated within a sample of the same ethnic background. Administering facial stimuli

manipulated along this dimension to non-Caucasian individuals would likely introduce substantial heterogeneity, particularly given the hypothesized role of familiarity in the mechanisms under investigation. For this reason, we restricted participant recruitment to Caucasian individuals. However, we are aware of the influence of cultural context in shaping social and occupational stereotypes; as well as its influence in shaping individual's support- seeking behaviours and responsiveness to social stimuli (Wu et al., 2021). Therefore, further studies should address this issue to clarify the potential role of cultural context in social pain modulation.

Finally, we are aware that presenting different identities for each condition, rather than a single identity, may have contributed to increasing the variability in the collected results. However, we emphasize that despite this potential variability, the fact that different facial stimuli manipulated to resemble the stereotypical representation of a "good doctor" consistently elicited an overall placebo hypoalgesic effect on pain perception—particularly concerning the affective component—not only corroborates the validity of our dimension as a valuable research tool but also strengthens the hypothesis that inferences from social judgments are sufficient to impact clinically relevant outcomes related to pain.

Moreover, we acknowledge that the "good doctor" models are based on group ratings and thus represent a generalized, stereotypical representation. While this approach allows us to examine the generalizability of social stereotypes, it also prompts further investigation, specifically into the effects of individualized, idiosyncratic models on social context and pain perception—a focus of our subsequent studies.

4.5 Conclusions

In conclusion, our findings demonstrate that faces embodying the "good doctor" stereotype–characterized by familiarity and competence, among other social attributes–can modulate pain perception, particularly its affective component, functioning as a safety signal that reduces the threat associated with the painful experience. This effect is further influenced by individual genetic predispositions related to socio-affiliative behavior and placebo hypoalgesia mechanisms. These results underscore the importance of social context, individual differences, and the role of familiarity in activating psycho-neurobiological mechanisms (i.e., placebo mechanisms) shaping pain experiences. These findings underscore the adaptive role of socially salient cues, such as faces, in shaping pain responses. Elucidating these mechanisms will contribute to clinical approaches that promote patients' recovery and overall health outcomes and provide a foundation for future research

aimed at disentangling these complex interactions, including investigations into individualized representations and cross-cultural variations.

5 Visualizing the idiosyncratic representation of a "Good doctor": development and validation

5.1 Introduction

In the previous study, we showed the modulatory effect that faces aligning with the stereotypical representation of a "good doctor" exert on the subjective experience of pain. We also highlighted the role of individual differences in shaping both the intensity and the direction of this modulation, with a particular focus on the effects of genetic variations related to neurotransmitter systems involved in pain modulation (i.e., dopaminergic and endocannabinoid systems) and socio-affective behaviours (i.e., serotoninergic and oxytocinergic systems).

Notably, individual differences may influence facial- based inferential processes as well (for more details, see subchapter 1.3). Indeed, the way facial features are weighted in the judgment process varies significantly across individuals; and although the evidence remains far from conclusive, this variability may be partially influenced by different factors, as gender, age or even their political ideology (Albohn et al., 2022, 2024, 2025; Hehman et al., 2017; Martinez et al., 2020). Moreover, when the social dimension under consideration becomes more complex and less directly tied to specific physical traits, as in the case with occupational roles like the "good doctor", individual impressions and personal biases tend to play a greater role than shared, universally recognizable facial features (Albohn et al., 2024). In other words, since the idea of a "good doctor" depends largely on personal stereotypes, people's judgments are likely to vary considerably. It follows that although reliable in representing the stereotypical representation of a "good doctor" at the group- level, the "good doctor" dimension we developed may not be the best approximation of this stereotype at the individual level.

Building on this perspective, the present study aimed to capture, quantify and visualize idiosyncratic representations reflecting the physical and social attributes that individuals associate with their stereotypical representation of a "good doctor". Using a data-driven method that leverages generative artificial intelligence (Albohn et al., 2025), we sought to generate and validate psychologically aligned models of the "good doctor" – that is, robust and photorealistic models reflecting participants' idiosyncratic representations of the stereotype. Furthermore, given the increasingly consistent evidence highlighting the crucial role of intrinsic familiarity in influencing relevant medical outcomes such as perceived pain (Bagnis et al., 2023; Mattarozzi et al., 2021), and considering the importance

this dimension may play in shaping the "good doctor" stereotype, as suggested by the previous results within this project- the present study intends to verify whether judgments of intrinsic familiarity also contribute to "good doctor" idiosyncratic representations.

To address these objectives, we conducted two experiments. Experiment 1 focused on the development of the "good doctor" idiosyncratic models. Participants made categorical judgments of a series of hyper-realistic AI-generated faces, sampled from a latent multidimensional space. These judgments were used to build "good doctor" idiosyncratic models, reflecting the idiosyncratic representation of a "good doctor"- as well as their counterpart- at the level of the individual participant.

Experiment 2 focused on the validation of these models. In a face rating task, participants judged the facial stimuli generated from their idiosyncratic models and the models of other participants on two social judgments, namely their resemblance to a "good doctor" and their perceived intrinsic familiarity. We predict that, although sharing some commonalities, the idiosyncratic "good doctor" models will appear quite diverse from each other. Consistently, we expect to find that facial stimuli manipulated along the "good doctor" idiosyncratic models would be rated according to the models' predictions (i.e., "high good doctor" faces will be rated as more closely resembling a "good doctor" compared to "low good doctor" faces). However, in agreement with recent findings emphasizing the pivotal role of idiosyncratic contributions in forming social judgments, we expected that idiosyncratic visual models built on individual participant judgments will be more effective in predicting their own ratings compared to the ratings of other participants. Finally, in line with the familiar face overgeneralization hypothesis (Zebrowitz et al., 2007)– which posits that individuals with familiar features are usually perceived as non- threatening and signaling safety– we predicted that "high good doctor" faces will be rated by participants as more familiar than "low good doctor" faces. We further predicted that participants would judge faces generated using their own idiosyncratic model as more familiar than those generated using other participants' models.

5.2 Experiment 1

5.2.1 Materials and Methods

5.2.1.1. Participants

107 participants (74 women, 30 men, 3 non-binary; age range= 18–70; M = 27.06; SD = 10.93) were recruited through the Virtual Lab of the University of Chicago Booth School of Business, USA.

Of these, 17 were excluded from the analyses due to low intra-rater correlation (r<.05)(N=11) or limited variability in responses to repeated stimuli, which prevented us from generating individual participant models (N=6).

Thus, the final sample for this experiment consisted of 90 participants (62 women, 25 men, 3 non-binary; age range = 18–70; M=26.81; SD=9.77), who self-identified as follows: Caucasian (N=40), Asian (N=33), African American (N=5), Hispanic (N=5), Latin American (N=3), Arabic (N=1), Indian American (N=1), Indo Caribbean (N=1) and Middle Eastern (N=1).

The sample size is consistent with previous studies that aimed to obtain stable estimates of idiosyncratic contributions in social judgment formation (i.e., approximately 60 participants evaluating 30 stimuli, each presented twice) (Albohn et al., 2025; Martinez et al., 2020).

The study was approved by the Aura Institutional Review Board of the University of Chicago and pre-registered on AsPredicted.org (https://aspredicted.org/207268).

5.2.1.2 Stimuli

300 facial stimuli (150 male, 150 female) were selected from a pool of 1,784 randomly generated faces using the Deep Learning model NAMFHQ-config-f-004000.pt (Albohn et al., 2025). Specifically, the selection of these stimuli was based on the following inclusion criteria: the faces had to appear as adults; they could not exhibit artifacts, shadows, or other elements that compromised image quality; the background had to be non-distracting, avoiding overly colorful or flashy elements; the face could not be obscured by distracting items such as sunglasses, hats, or prominent jewelry; and the face had to be oriented toward the centre of the screen.

The stimuli were presented via a Qualtrics link. 30 randomly selected stimuli were presented twice, in order to allow for the calculation of intra-rater reliability.

Therefore, the total number of presented stimuli was 330.

5.2.1.3 General procedure

Six participants per day were recruited over a period of 18 days to take part in a two-phase study on “judgments of faces”, conducted through the Virtual Lab webpage of the University of Chicago Booth School of Business (Chicago, IL). The limited number of participants per day was established based on the procedure described in Albohn et al., 2025, to ensure that each participant would evaluate only

a subset of the images generated from the idiosyncratic models during the validation phase and to facilitate task administration. Hence, 6 participation slots were made available each day. After reading the study description and signing the informed consent, participants could reserve one of the available slots by clicking on the corresponding link. By doing so, they committed to completing the first of the two parts of the study by 5:00 PM on the same day, thereby preventing other users from accessing the same link.

If a participant failed to complete the study by the deadline, the slot was reopened, allowing new participants to enroll until the target of six completed sessions per day was met.

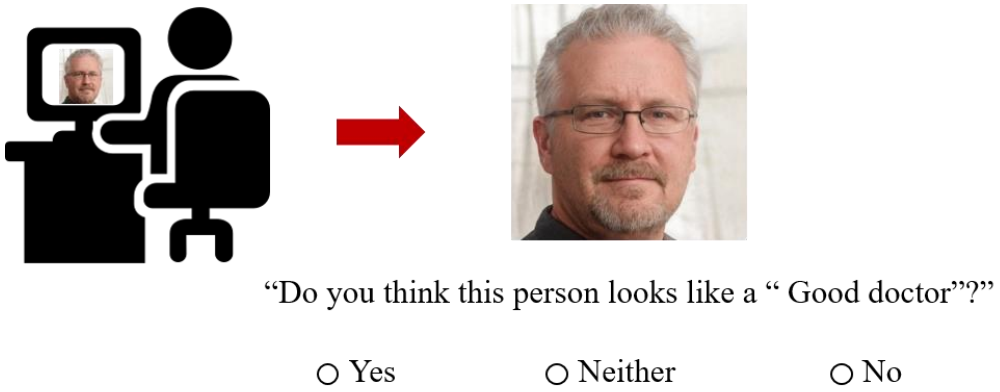
Upon accessing the link and signing the informed consent form, participants were instructed to indicate, based on their first impression, whether the presented facial stimuli resembled a "good doctor" by choosing among three response options: "Yes", "No", "Neither".

Two practice trials were administered to help participants familiarize themselves with the procedure. Then, face stimuli were displayed sequentially on the screen with the question "Do you think this person looks like a "good doctor"?".

An example trial is illustrated in Figure 5.1.

Figure 5.1

Example trial of the task administered during Experiment 1



Upon completion of the task, participants received an email containing the link to the second experiment, which became active on the following day.

Participants were informed that they had 48 hours to complete both experiments; and that, once completed, they would receive $12.

5.2.1.4 Idiosyncratic model construction

For the construction of the idiosyncratic models, we followed the procedure described in Albohn et al., 2025. Each of the presented photorealistic facial stimuli was defined as a matrix of numeric values in the latent space. For each participant, we calculated the average of all the stimulus latents judged as not resembling a "good doctor" and we subtracted it from the average of all the stimulus latents judged as resembling a "good doctor". Then, we added back the average of all the stimulus latents categorized as "Neither". This yielded a directional vector in the model's latent space. By linearly interpolating along this vector (i.e., ± 1 SD, ± 0.5 SD and 0), we generated visualizations of each participant's idiosyncratic model at varying intensities. Specifically, we created five images for each participant's model: two "high good doctor" faces (i.e., faces representing the linear interpolation levels at +0.5 and +1), characterized by the physical features the participant associates with their stereotypical image of a "good doctor"; two "low good doctor" faces (i.e., faces representing the linear interpolation levels at -0.5 and -1), whose physical features are very distant from the participant's stereotypical image of a "good doctor"; and one neutral face, reflecting the average of the latents of all images they categorized as "neither" during the experiment.

5.2.1.5 Data analyses

To assess test- retest reliability for each participant, we computed the correlation between the ratings participants gave to repeated stimuli. Participants with correlations less than or equal to .05 were excluded from subsequent analyses (Albohn et al., 2025).

Moreover, to assess the similarity between participants' idiosyncratic visual models, we computed cosine similarities between the difference score vectors reflecting each participant's model—namely, the vectors obtained by subtracting the average of all stimulus latents judged as not resembling a 'good doctor' from the average of all stimulus latents judged as resembling a 'good doctor', and then adding back the average of all stimulus latents categorized as 'Neither'. Cosine similarity (CS) reflects the similarity of two vectors in the latent space. Values closer to 1 indicate that the models are highly similar, whereas values closer to 0 and -1 indicate, respectively, no similarity and anti-similarity (Levy et al., 2025). Mean cosine similarity (M_{cs}) between all models was then calculated.

The alpha level for all analyses was set to $p<.05$.

Analyses were performed using RStudio (version 4.4.1; 2024-06-14 ucrt, Posit, PBC, Boston, MA, USA).

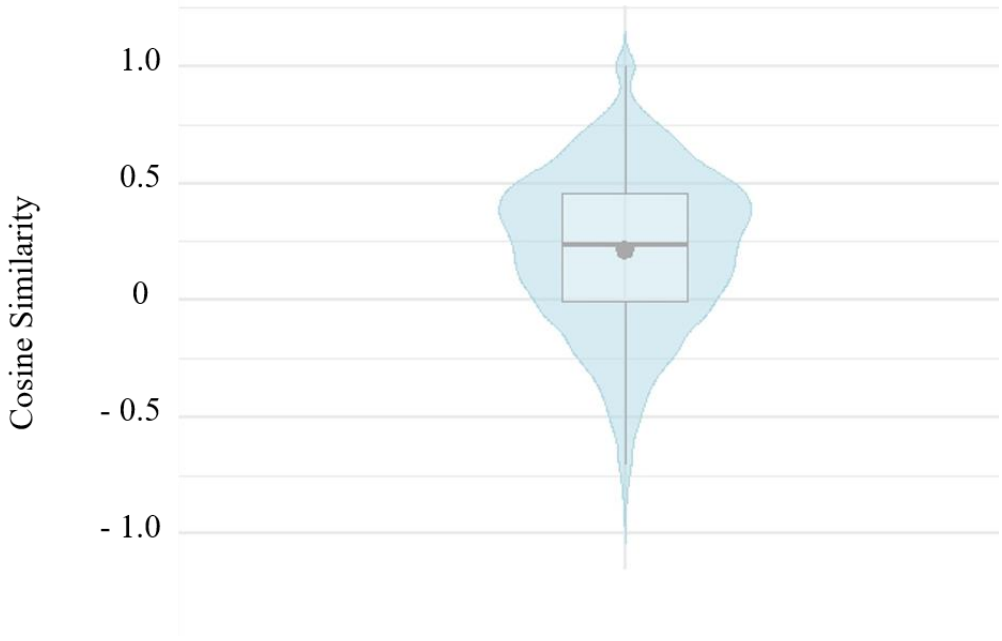
5.2.2 Results

Mean intra-rater reliability coefficient (r=.55) indicated moderate consistency across participants, suggesting that the collected responses reliably reflect their individual judgments.

Furthermore, mean CS (M_{cs}=.21) evidenced that although sharing some similarities, participants' idiosyncratic models were quite diverse to each other. The distribution of the CS CSs between each participant's model and the models of the other participants is displayed in Figure 5.2.

Figure 5.2

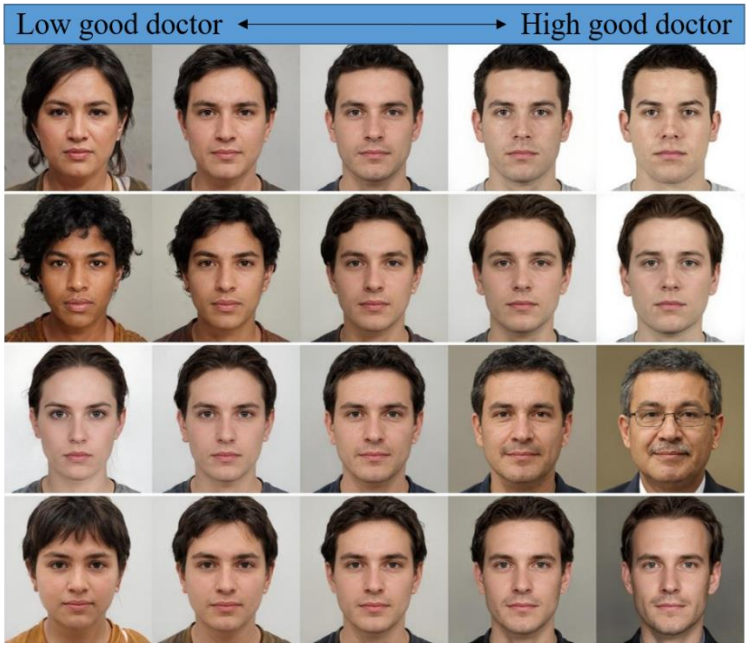
Distribution of pairwise cosine similarity values between participants' idiosyncratic models in Study 3. Individual data points are shown using a quasirandom dot plot to avoid overlap. The black horizontal line indicates the mean cosine similarity, while the blue horizontal line indicates the median



Examples of "good doctor" idiosyncratic models are displayed in Figure 5.3.

Figure 5.3

Examples of "good doctor" idiosyncratic models.



Note. Each row displays images reflecting participants' idiosyncratic visual models of the "good doctor" stereotype. The central image of each row represents the average of the latents participants judged as "neutral". Images to the left and the right reflect the linear interpolation at -1, -0.5 and 1, 0.5; which represent, respectively, the "low good doctor" and the "high good doctor" idiosyncratic faces.

5.3 Experiment 2

Experiment 1 aimed to develop idiosyncratic models of the "good doctor" stereotype, reflecting the participant's individual representation of what a doctor should look like to be perceived as "good" at first glance.

By linearly interpolating along the vector at + 1 SD/0.5 SD, -1 SD/-0.5 SD and 0, these models allowed us to visualize the facial features each participant associates with a "good doctor", as well as those features they consider as least representative of this stereotypical representation.

Although the models shared some commonalities– since they all refer to the general stereotype of a "good doctor" – analyses of the mean cosine similarity showed that they were nonetheless quite distinct from one another. Hence, validation was required to assess whether these representations effectively reflected this stereotype at the individual level.

In Experiment 2, we asked participants who had completed Experiment 1 to judge both idiosyncratic and non- idiosyncratic representations of the "good doctor" stereotype in terms of their perceived resemblance to a "good doctor". This procedure allowed us to test the predictive validity of the models, observing whether the idiosyncratic representation of a "good doctor" would be more likely to be judged by each participant as more similar to this stereotype compared to non- idiosyncratic representation.

Furthermore, consistent with the role that familiarity seems to play in influencing relevant clinical outcomes (Bagnis et al., 2023; Mattarozzi et al., 2021), and considering the significant correlation between this dimension and the group-level, "good doctor" dimension developed in Study 1 (for more details, see Chapter 3), we also asked participants to evaluate idiosyncratic and non-idiosyncratic representations along this dimension. We expected to find that idiosyncratic models would be more accurate in predicting familiarity ratings of participants than non- idiosyncratic models, coherently with the hypothesis according to which the "good doctor" stereotype may appear as highly familiar-looking.

5.3.1 Materials and Methods

5.3.1.1 Participants

Of the 90 participants included in the statistical analyses of Experiment 1, only 1 participant did not complete Experiment 2. Four participants were excluded from further analyses for the following exclusion criteria: low intra-rater correlation ($r<.05$) (N=1); they were identified as outliers for the dependent variables resemblance with a "good doctor" and "familiarity" ratings (i.e., their Z scores for these variables exceeded ±3) (N=2); error in the study administration (i.e. participant did not rate images from their idiosyncratic model).

The final sample for this experiment consisted of 85 participants (59 women, 23 men, 3 non-binary; age range = 18–70; M=26.87; SD=9.91), who self-identified as follows: Caucasian (N=39), Asian (N=32), African American (N=5), Hispanic (N=5), Latin American (N=2), Indian American (N=1) and Middle Eastern (N=1).

The sample size is consistent with previous studies that aimed to obtain stable estimates of idiosyncratic contributions in social judgment formation (i.e., approximately 60 participants evaluating 30 stimuli, each presented twice) (Albohn et al., 2025; Martinez et al., 2020).

The study was approved by the Aura Institutional Review Board of the University of Chicago, USA.

5.3.1.2 Stimuli

The Deep Learning model NAMFHQ-config-f-004000.pt (Albohn et al., 2025) used in Experiment 1 was implemented to generate "good doctor" idiosyncratic models based on the judgments collected through the previous experiment. Specifically, we developed 5 different images for each participant: two "high good doctor" faces (i.e., faces representing the linear interpolation levels at +0.5 and +1), reflecting participant's stereotype of a "good doctor"; two "low good doctor" faces (i.e., faces representing the linear interpolation levels at -0.5 and -1), whose physical features are very distant from this stereotype; and, finally, one neutral face.

The total number of stimuli generated for this second phase of the study was 525 stimuli (i.e., 5 images for the 105 participants whose responses, collected during Experiment 1, enabled the generation of the idiosyncratic models), administered through a Qualtrics link. Notably, each participant was presented with the facial stimuli generated using their idiosyncratic model and those generated from the models of 5 other participants; for a total of 30 facial stimuli (5 faces per model x 6 idiosyncratic models = 30 facial stimuli) randomly presented and repeated across three different blocks. In the first two blocks, the faces were rated based on their similarity to a "good doctor"; in the third block, they were judged in terms of perceived familiarity.

Therefore, participants rated 90 facial stimuli in total (30 stimuli repeated in 3 blocks = 90 facial stimuli).

5.3.1.3 General procedure

For this second experiment, participants were invited to click on the link they received the previous day. Once they signed the informed consent, they were instructed to indicate, based on their first impression and using a numeric scale from 0 ("Not at all") to 7 ("Extremely"), to what extent the displayed facial stimuli looked like a "good doctor" or "familiar".

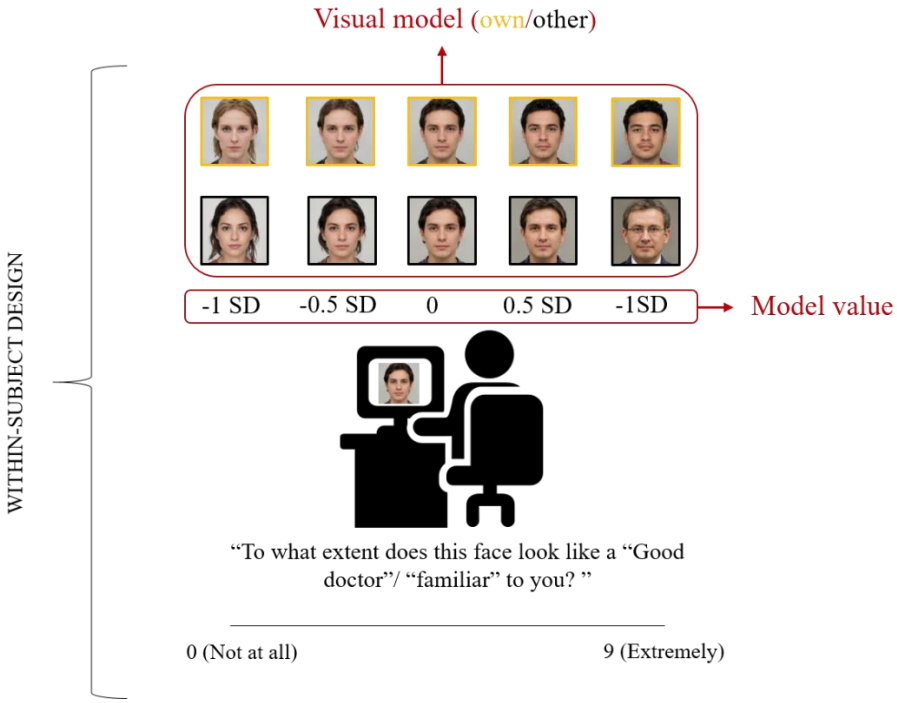
Two practice trials were administered to help participants familiarize themselves with the procedure. Then, the facial stimuli previously selected were displayed sequentially on the screen and repeated across three different blocks. In the first block, facial stimuli were paired with the question "To what extent does this face look like a "good doctor" to you?". This block was repeated twice, to assess intra-rater reliability on the main dependent variable. In the third block, instead, they were followed by the question "To what extent does this face look "familiar" to you?"

Notably, the stimuli presented corresponded to those generated through the idiosyncratic models of the six participants recruited on that specific day. This procedure ensured that each participant could rate both the 5 images generated by their own model and the 25 images generated from the models of the other 5 participants.

An example trial is illustrated in Figure 5.4.

Figure 5.4

Example trial of the task administered during Experiment 2



Experiment lasted about 15 minutes. Once they finished the task, participants were redirected to a webpage for compensation.

5.3.1.4 Data analyses

To assess the intra-rater reliability, we computed the correlation between the ratings participants gave to repeated stimuli. Participants with correlations less than or equal to 0.05 were excluded from subsequent analyses (Albohn et al., 2025).

Linear mixed-effects regression models were used to analyze the effect of the two independent variables– model value (levels of linear interpolation at ±1/±0.5/0) and visual model ("own"/ "other")

– on the ratings for each of the two dependent variables considered, namely resemblance with a "good doctor" and familiarity. Notably, model value variable was used to classify the generated images into "high good doctor" images (with a model value of +1 and +0.5), "low good doctor" images (with a model value of -1 and -0.5), and neutral images (with a model value of 0); while visual model variable indicated whether the images were generated using the participant's idiosyncratic model or the models of other participants.

Model value and visual model were considered as fixed effects in the model, while participants and stimuli were included as random effects, with random intercepts only.

The significance level was set at $p<0.05$ for all analyses.

Data analysis was conducted using RStudio (version 4.4.1; 2024-06-14 ucrt, Posit, PBC, Boston, MA, USA).

5.3.2 Results

Mean intra-rater reliability coefficient ($r=0.67$) indicated moderate consistency across participants, suggesting that the collected responses reliably reflect their individual judgments.

The regression model analysis revealed a significant positive effect of model value on the dependent variable "resemblance with a good doctor" ($b=1.54$, $t(2274.54) =18.64$, $p<.001$), showing that facial stimuli at higher levels of linear interpolation were rated as more similar to the face of a "good doctor".

A significant interaction between model value and visual model was also observed ($b=-0.99$, $t(2046.81) =-12.04$, $p<.001$), indicating a stronger influence of the model value when the face was generated using participant's own idiosyncratic model compared to that of another participant.

No significant effect was found for the visual model variable considered independently ($b=0.02$, $t(2046.81) =0.40$, $p=.71$).

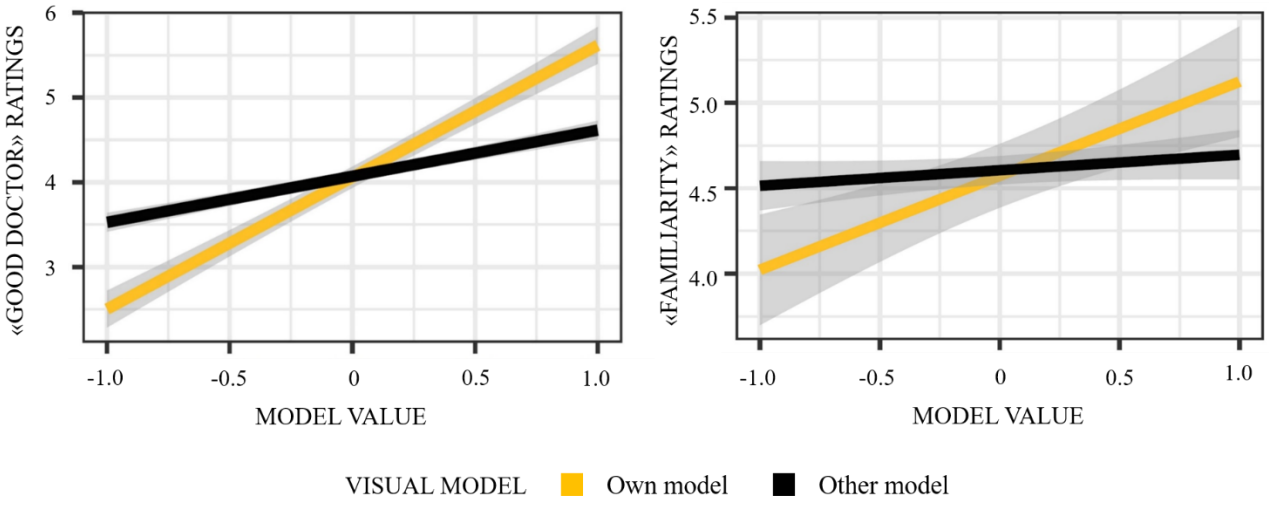
Similar results were found for the dependent variable "familiarity." Specifically, it was revealed a positive effect of model value on familiarity ratings ($b=0.55$, $t(2455.64) =5.31$, $p<.001$), as well as a significant interaction between model value and visual model ($b=-0.45$, $t(2077.02)=-4.07$, $p<.001$) evidencing a stronger influence of model value on familiarity ratings when the face was generated using participant's idiosyncratic model rather than another participant's model.

Finally, no significant effect was found for the visual model variable considered independently (*b*=0.03, *t*(2077.02)=0.39, *p*=.70).

Results of the effects of model value and visual model on the resemblance with a "good doctor" and familiarity ratings are displayed in Figure 5.5.

Figure 5.5

Validation results of the effects of model value (levels of linear interpolation at ±1/±0.5/0) and visual model (own vs. other) on the resemblance with a "good doctor" and "familiarity" ratings.



Note. Each participant judged images generated from their own visual model (yellow regression lines) and those generated from a subsample of other participants' visual models (black regression lines). The x-axis represents the model interpolation values and the y-axis the participants' responses. Shaded areas around each line display 95% confidence intervals.

5.4 Discussion

When making social judgments from others' faces, the complexity of the social dimension under consideration matters significantly. As the connection between that dimension and specific physical traits becomes less direct– as in the case of occupational roles such as the "good doctor" stereotype– individual impressions and biases exert a stronger influence on face-based social judgments (Albohn et al., 2024). This highlights the need to capture, quantify, and visualize idiosyncratic representations reflecting the physical and social attributes that individuals associate with their own stereotypical representation of a "good doctor".

Through the use of generative models, we developed individual-level, idiosyncratic representations of the "good doctor" stereotype – that is, visual models reflecting the physical features and the social attributes that each participant associates with their own stereotypical representation of a "good doctor". In line with previous research highlighting the relevance that idiosyncratic contributions have in shaping complex social judgments (Albohn et al., 2025; Martinez et al., 2020), we found that participants' own models were more accurate in reflecting their stereotypical representation of a "good doctor" compared to the models of other participants. Moreover, consistent with past research highlighting the role of intrinsic familiarity in modulating relevant medical outcomes, such as pain perception and treatment satisfaction (Bagnis et al., 2023; Mattarozzi et al., 2021), as well as the results obtained in previous studies within this thesis, our results point toward this dimension as pivotal in building this stereotype.

Mean cosine similarity, a measure reflecting the similarities between vectors in the latent space (Mana & Sasipraba, 2021), among the generated idiosyncratic models showed that although sharing some commonalities, the "good doctor" idiosyncratic models were diverse across participants. This result aligns with our expectations: unlike judgments as masculinity/femininity and age - which tend to be reliably mapped onto facial features and show higher observer agreement (Albohn et al., 2025; Mogilski & Welling, 2018) – complex judgments, as familiarity for instance (Albohn et al., 2025), involve less consensus regarding which features should be considered and how they should be weighted (Albohn et al., 2024, 2025). This suggests that although individuals may agree on the social traits defining occupational stereotypes and social group memberships (Bin Meshar et al., 2022), they may exhibit variability in the physical features they associate with these social attributes.

This interpretation is further supported by results obtained in the validation study. We found that participants' own idiosyncratic models were more accurate in predicting their judgments than other participants' models. In other words, whereas "high good doctor" faces were generally perceived as more similar to the stereotype, confirming the existence of shared features, those generated from participants' own models aligned more closely with their individual representations. These findings corroborate and extend previous research (Albohn et al.,2025; Martinez et al., 2020), showing that idiosyncratic contributions outweigh shared ones not only in personality trait judgments, but also in social class and occupational stereotypes.

Similar results were found in relation to judgments of familiarity. "High good doctor" faces were generally rated as more familiar than "low good doctor" faces, regardless of the model used to generate them. However, participants' own idiosyncratic models were more accurate in predicting

their familiarity judgments than other models, with idiosyncratic “high good doctor” faces rated as more familiar than the others.

The implications of these findings are twofold. On one side, they strengthen the evidence for the idiosyncratic nature of the “good doctor” stereotype, since familiarity emerged as a strong idiosyncratic attribute (Albohn et al., 2025). On the other, they reinforce the key role of intrinsic familiarity in shaping patients' impressions of their healthcare providers. Specifically, high familiar-looking people may act as “safety signals” (Debiec & Sullivan, 2017), triggering comforting associations that boost one's sense of security in facing potential threats. In healthcare, these may reduce patient’s negative feelings and improve clinical outcomes (Bagnis et al., 2023; Mattarozzi et al., 2021); making this feature particularly desirable in a physician and transforming it into a pivotal dimension in shaping the “good doctor” stereotype.

Finally, exploratory analyses on the idiosyncratic models participants generated indicated that the “good doctor” stereotype predominantly features male doctors (for more details, see the subchapters 8.2.1 and 8.2.2 in the Supplementary Materials). In other words, despite the success in medical training and clinical practice women doctors have achieved (Tsugawa et al., 2017; Wallis et al., 2017), as well as their increasing representation in the physician workforce (Butkus et al., 2018), both male and female participants were inclined to associate the “good doctor” stereotype with traditional male figures. These sex-related occupational stereotypes persist across various professional domains. For instance, although women’s representation in STEM disciplines is growing worldwide (Murphy & Oesch, 2016), disparities are still evident (Mcquillan & Hernandez, 2021), with male STEM managers often displaying implicit in-group gender favouritism in their hiring decisions (Friedmann & Efrat-Treister, 2023). Moreover, fields as military and aviation remain predominantly male work environments and are highly masculinized, where women face harassment and workplace discrimination (Bridges et al., 2020), as well as social exclusion and marginalization (Burgess et al., 2021). Finally, in healthcare, female doctors still struggle to be recognized in their professional role by their patients (Haynes et al., 2021; Stavely et al., 2022), evidencing how people continue to see the “good doctor” as a male.

Consistently, our results align with previous findings highlighting persistent sex-related biases towards female physicians (Haynes et al., 2021; Kauff et al., 2022; Stavely et al., 2022). Importantly, our results also evidenced that this pattern was not mirrored in the generation of “low good doctor” faces, which instead did not appear to be associated with a specific biological sex.

In addition to sex biases, age-related perceptions also emerged. Specifically, the "good doctor" stereotype was perceived as older compared to the "low good doctor" stereotype, which aligns with the common perception that older doctors are seen as more competent (Hall et al., 2020). While this contrast with findings that older people are generally perceived as warmer and interpersonally skilled, but incompetent (Cuddy, 2004; Gordon & Arvey, 2004; Redman & Snape, 2002), such generalization may not apply to high- skill professions like medicine. Moreover, the "high good doctor" faces developed in this study did not represent elderly doctors per se, but rather those who appeared older than their counterparts. Thus, more "mature" doctors, rather than elderly ones, may be perceived as reflecting greater competence and experience in patient care.

The present research has some limitations worth noting. First, the facial stimuli used to generate idiosyncratic models were not balanced in terms of sociodemographic attributes such as race and age. This may have exerted an influence on participants' idiosyncratic representations of the "good doctor," potentially biasing them toward specific physical characteristics. In this regard, it is noteworthy to mention that the administered facial stimuli were hyper-realistic faces generated artificially; therefore, their sociodemographic features emerged purely through perceptual inferences, without explicit labelling. Moreover, the primary aim of the study was to investigate the effectiveness of idiosyncratic models in reflecting participants' representation of a "good doctor", rather than to analyse the specific features of such images per se. Future research should investigate this aspect more explicitly —for instance, by observing whether individuals from different racial backgrounds are more likely to prefer doctors perceived as belonging to their own group versus others.

Second, most of our participants were women, which may have partially contributed to the similarity observed across the idiosyncratic models. Nevertheless, our results did not reveal sex- related differences in generating males and females idiosyncratic "high good doctor" faces, supporting the notion that a bias toward the "good doctor" stereotype as a male appear equally robust in males and females. Interestingly, our findings are in contrast with previous research indicating that women, more than men, tend to dissociate dominant, science-based disciplines from male figures (Koch et al., 2015; March et al., 2016); highlighting that the observed similarities cannot be solely attributed to a higher proportion of women in our sample. Future studies should therefore explore this issue more in depth, investigating whether gender and other socio-demographic factors may contribute to shape the individual stereotypical representation of different occupational roles, including doctors.

5.5 Conclusions

First impressions are crucial in influencing individuals' behaviours and attitudes, even when they do not accurately reflect the true attributes of others (Zebrowitz & Montepare, 2008). In making such judgments, however, individuals differ considerably in how they interpret and weight facial features (Albohn et al., 2025).

In the healthcare context, these face-based inferences have been shown to significantly impact medical outcomes (Bagnis et al., 2023; Hall et al., 2020; Necka et al., 2022). Since first encounters often lay the groundwork for the patient–physician relationship, understanding the mechanisms behind such impressions becomes increasingly relevant, especially in this field.

In the present study, we provide evidence that individuals hold highly personalized mental representations of the "good doctor" stereotype, namely of what they think a valid doctor would look like. These representations are shaped by both shared stereotypes and idiosyncratic preferences; and despite some commonalities, they uniquely reflect the stereotypical representation that participants hold in their minds. Notably, our findings suggest that these idiosyncratic models are also strongly correlated with intrinsic familiarity.

Integrating these personalized models, particularly considering the role of familiarity and demographic factors, into future experimental paradigms may open promising avenues for advancing patient-centered care, improving the quality of treatments provided and paving the way for more personalized medicine.

6 The effects of idiosyncratic “good doctor” stereotypical representation on pain perception

6.1 Introduction

Study 2 revealed that faces manipulated along the “good doctor” dimension– to incorporate the physical traits typically associated with this stereotype–appeared to modulate pain perception. Nonetheless, face manipulation accounted for only a small proportion of the models’ variance; suggesting that a larger proportion of variability was attributable to intra-rater differences. Consequently, participants may have experienced pain modulation to different degrees, depending on the extent to which the presented faces aligned with their own mental representation of a “good doctor”.

Results from the Study 3 clearly indicated that individuals differ in the way they visualize their own stereotypical representation of a “good doctor”. While sharing certain common features, the idiosyncratic models generated were diverse from one another, and proved to effectively capture how a “good doctor” appears in each individual’s mental representation.

Building on this perspective, we have implemented a new study aiming to observe whether idiosyncratic representations of the “good doctor” stereotype would modulate pain perception more effectively than non-idiosyncratic representations. Leveraging the generative models described in the previous study, we create idiosyncratic models for each participant recruited. These models are then used to generate idiosyncratic “high good doctor” and “low good doctor” faces, which are administered to participants while undergoing painful stimulations. Notably, each participant is presented with faces generated using their own models and the model of another participant, to test whether differences in pain modulation could be observed across the diverse conditions.

Specifically, the study consists of two phases.

The first phase focuses on the construction of idiosyncratic models. Following the same procedure described in the previous chapter (for more details, see subchapter 5.2.1), we generate each participant’s “good doctor” idiosyncratic model.

The second phase consists of the administration of a pain task and a familiarity rating task.

The pain task is designed following a procedure similar to that implemented in Study 2 (for more details, see subchapter 4.2). Participants are exposed to facial images derived from their own "good doctor" idiosyncratic model and from another participant's model while receiving painful electrical stimuli. Pain intensity and unpleasantness are measured after each painful stimulation, and HR and HRV components (i.e., RMSSD, VLF, LF, HF, HF/LF) are continuously recorded throughout the task, as physiological indicators of sympathetic and parasympathetic activation and perceived pain (Lavigne et al., 2001; Tousignant-Laflamme et al., 2005; Treister et al., 2012).

For the familiarity rating task, instead, faces presented during the pain task and new faces are administered to participants, who are asked to judge them in terms of perceived familiarity.

We predict that faces reflecting participant's idiosyncratic representation of a "good doctor" would be associated with stronger pain modulation compared to non-idiosyncratic faces. As previously suggested, this effect is expected to be mediated by inferences of "familiarity". Indeed, "high good doctor" idiosyncratic faces have been perceived as not only reflecting how a doctor should look like to be perceived as "good" on an individual basis, but also as more familiar than non- idiosyncratic models. Therefore, we anticipate that these faces may act as particularly effective "safety signals", reducing the perceived threat and thereby leading to a more pronounced attenuation of pain. In agreement with this hypothesis, we also expect to find that "high good doctor" idiosyncratic faces would be judged as more familiar-looking than the other faces, irrespective of whether these faces have been already seen or not.

6.2 Materials and Methods

6.2.1 Participants

60 participants (30 women, 30 men, range age: 18-65 years, M= 27.07, SD= 9.20) have been recruited to participate in a study on pain modulation. Exclusion criteria are applied as follows: i) medical/psychiatric conditions affecting pain perception (N=6); ii) failure to generate the idiosyncratic model (N=7); iii) incomplete participation in the second experimental part (N=6); iv) technical problems during task administration (N=5); v) Z-scores > ±3 on control and dependent variables (N= 4); vi) low intra-rater correlation (r<.05) (N=2).

Therefore, the current sample consists of 30 participants (15 women, 15 men; age range= 19–48; M=26.53, SD=7.56).

The target sample size was established through an a-priori calculation using G*Power software (Faul et al., 2007), which indicated that assuming a small effect size (f = 0.10), a correlation of 0.50 among repeated measures across 5 measurements, a minimum of 121 participants would be required to achieve a statistical power of 0.80 for alpha=0.05. Participant recruitment is ongoing in order to reach this target sample size.

The study was reviewed and approved by the Bioethics Committee of the University of Bologna (Prot. No. 0388067)

6.2.2 Stimuli

For the construction of the idiosyncratic models, we follow the procedure described in Albohn et al., 2025 and illustrated previously (for more details, see paragraph 5.2.1.4). 300 facial stimuli (150 males, 150 females) are administered through a link via Qualtrics. To assess intra-rater reliability, 30 stimuli are presented twice, resulting in a total of 330 facial images per participant (300 stimuli + 30 repeated).

The collected judgments are used to generate the vector reflecting each participant's "good doctor" idiosyncratic model. Each model is used to generate five different images, namely two "high good doctor" faces, two "low good doctor" faces and one neutral face; for a total of 265 facial images (5 images for 53 participants whose judgments, collected during the first part of the study, allowed the generation of the idiosyncratic model).

For the pain task implementation, we select "high good doctor" (i.e., +1 interpolation level) and "low good doctor" (i.e., -1 interpolation level) faces from each participant's idiosyncratic model and the model of another participant, identified as the most dissimilar among all the models developed. Similarity between models is quantified using cosine similarity, which measures the similarity between the vectors representing the models in the latent space (Han et al., 2012). Lower cosine similarity values (closer to zero) indicate greater dissimilarity between two models.

Thus, four images per task version are selected for each participant. It derives that each participant completes a customized version of the task.

Finally, for the familiarity task, we select the "high" and "low good doctor" faces from each participant's model and three other models. Specifically, we included stimuli generated using the two most dissimilar models (lowest cosine similarity values) and the most similar model (highest cosine

similarity value). This results in eight stimuli per participant: four faces already seen in the pain task and four novel faces.

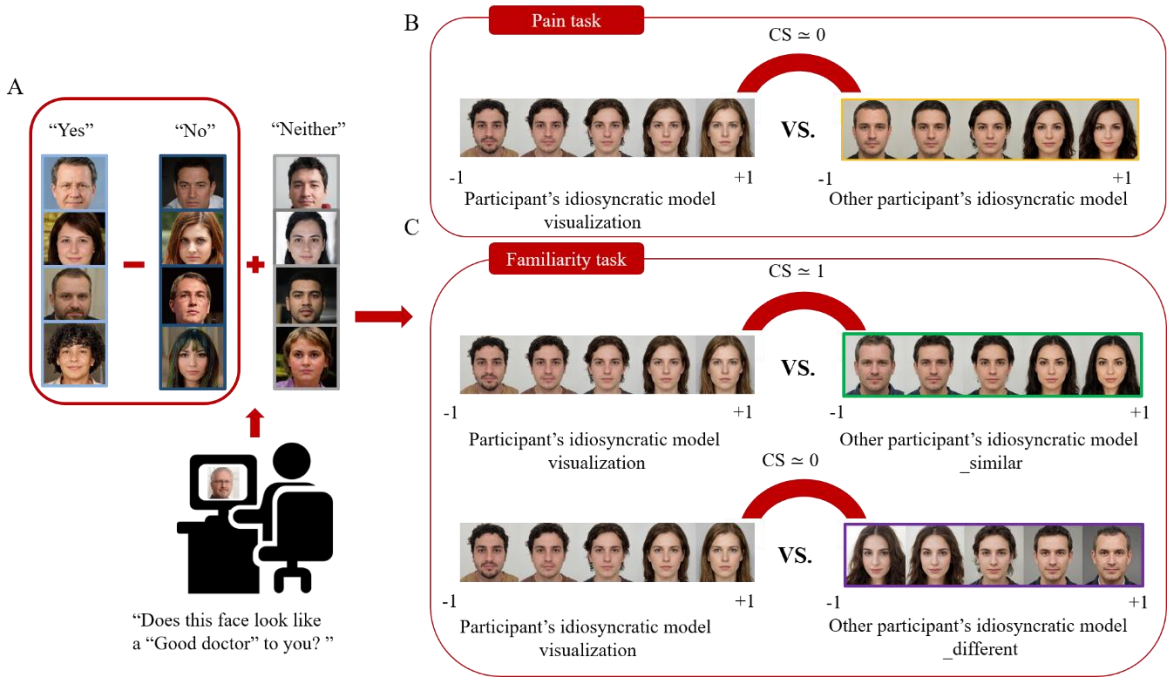
Stimuli presentation occurs via Qualtrics, and is randomized across participants.

Importantly, each participant's model is included once per condition (idiosyncratic, other, dissimilar, similar), ensuring that every participant completes a unique version of both tasks.

The procedure for the selection of the stimuli for both the pain and familiarity task is depicted in Figure 6.1

Figure 6.1

Procedure for the selection of the stimuli for both the pain and familiarity task.



Note. A) Creation of the idiosyncratic "good doctor" faces. For each participant, we calculated the average of all the stimulus latents judged as not resembling a "good doctor" and we subtracted it from the average of all the stimulus latents judged as resembling a "good doctor". Then, we added back the average of all the stimulus latents categorized as "neither". This yielded a directional vector in the model's latent space. By linearly interpolating along this vector (i.e., ± 1 SD, ± 0.5 SD and 0), we generated visualizations of each participant's idiosyncratic model at varying intensities. B) Selection of the other participant's idiosyncratic "good doctor" model, based on CS≃0. C)

Selection of the other two models generated using the ratings of other two participants. These models must be the most similar (CS≃1) and the second most dissimilar models among those generated.

6.2.3 General procedure

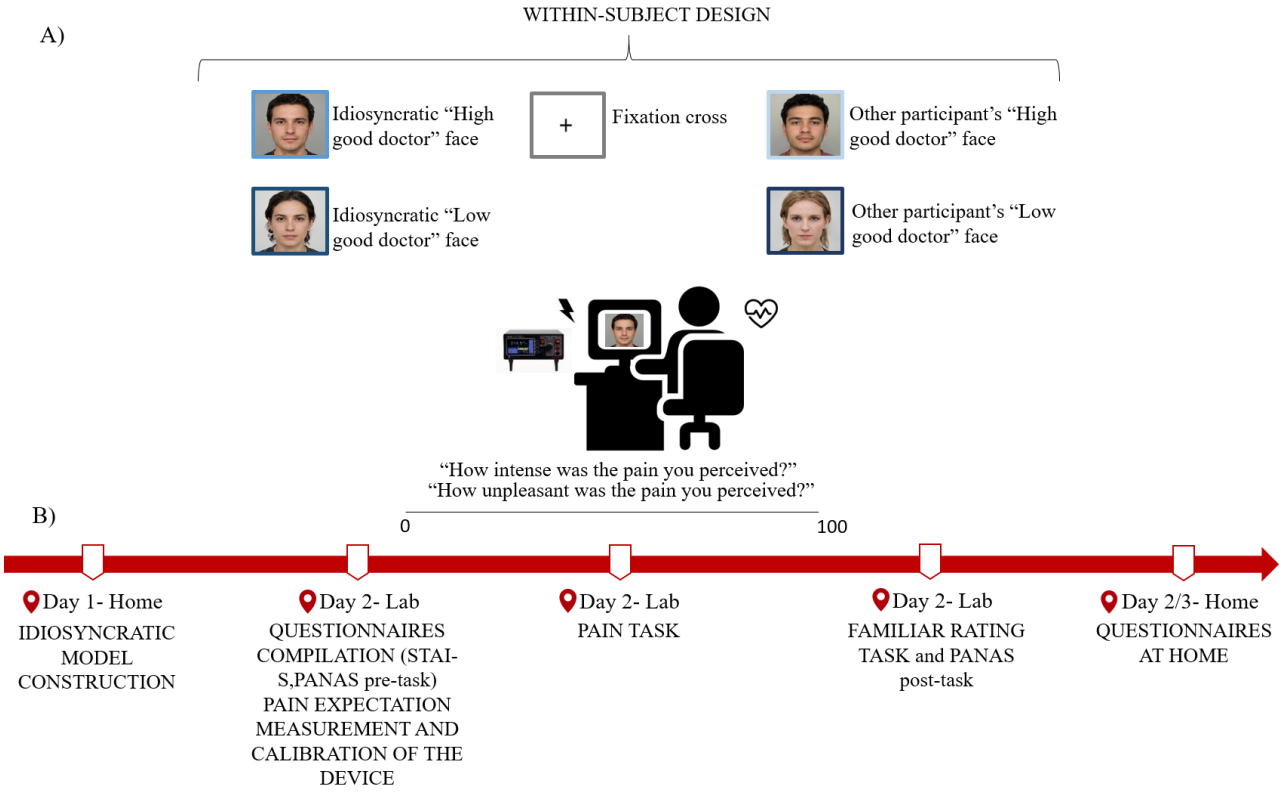
The experimental procedure consists of three phases, administered over three different days.

First, participants complete an on-line task for the idiosyncratic model construction. Second, they perform a pain task in which facial stimuli are paired with painful stimulations (i.e., electric impulses), followed by a familiarity rating task. Finally, participants complete additional questionnaires at home.

The overall experimental procedure is illustrated in Figure 6.2.

Figure 6.2

Pain task design (A) and the timeline of the experimental procedure (B).



Note. Panel A displays the pain task design, while panel B the timeline of the experimental procedure. Note that the order of block conditions (idiosyncratic "high/low good doctor", other participant's "high/low good doctor", fixation cross) is randomized across participants.

Phase 1: idiosyncratic model construction. After providing informed consent, participants complete a checklist to assess their eligibility for the study (e.g., absence of medical or psychiatric conditions that could affect pain perception). Then, they are sent an e-mail with a link to access the online task. Participants are hence instructed to view each facial stimulus presented on the screen and, based on their first impression, to indicate whether the face resembled a "good doctor" by choosing one of three options: "Yes," "No," or "Neither."

To allow participants to familiarize themselves with the procedure, two practice trials are administered before the experimental trials. Subsequently, facial stimuli are displayed sequentially on the screen, each paired with the question: "Do you think this person looks like a good doctor to you?"

Upon completion of the task, and once verified that the collected data allowed for model generation, participants receive a link to schedule an in-person appointment for the second phase of the experiment.

Phase 2: pain task and familiarity rating task administration. On the day prior to the experimental session, participants receive a reminder e-mail. Upon arrival at the laboratory, participants are asked to take a seat in a room where they first answer questions regarding exclusion criteria (health status, presence of medical conditions, painkiller intake 24h prior the experimental session) and then complete two self- reported questionnaires: i) the Positive Affect and Negative Affect Schedule (PANAS; Watson et al., 1988), aimed at measuring participants' positive and negative affective states and possibly indicating the individual's tendency to engage in avoid/approach behaviours in response to pain and environmental stimuli (Day et al., 2022); and ii) State- Trait Anxiety Inventory (STAI-S; (Spielberger, 1989), aimed at assessing participants' state anxiety, given its influence on pain perception (Frot et al., 2004; Gondo et al., 2012; Tang & Gibson, 2005).

Then, participants are invited to sit in front of a PC screen (screen size: 30 × 37 cm; visual angle: 2.78° × 4.28°; resolution: 1280 × 1024). Once comfortably seated upright, two electrocardiography electrodes are applied on the right clavicle and the lowest left rib for HR recording. HR is recorded

as beat-to-beat intervals in ms with the Bodyguard 2 (Firstbeat Technologies Ltd.; Parak & Korhonen, 2013).

Afterwards, they are informed about the subsequent procedure, which included i) the measurement of pain expectation and calibration of the electrical stimulation device (DS8R Bipolar Constant Current Stimulator, range 0–1000 mA), ii) the pain task, and iii) a familiarity rating task. Participants are also reminded that they can withdraw from the study at any time.

i) Pain expectation measurement and calibration of the device. The procedure follows that described in Study 2. Participant's pain expectation (i.e., "How much pain do you expect to feel during the procedure?") is measured using a 0 ("Not at all") -10 ("Extremely") numeric scale, and instructions about the tasks are provided. Two electrodes connected to the DS8R are attached to the participant's non-dominant forearm, and the instructions for determining tactile and pain thresholds are explained. The stimulus intensity used during the experimental task is set individually at twice the participant's pain threshold (Brączyk & Bąbel, 2021), and remains constant throughout the task.

ii) Pain Task. Following calibration, participants complete the pain task. 8 testing trials (6 trials without electric impulses + 2 trials with the electric impulses) are administered to allow participants to familiarize themselves with the procedure. Then, participants receive 30 electrical stimuli (6 electrical stimuli per 5 conditions, each lasting 100ms, identical in intensity, set at twice the pain threshold) paired with 5 visual stimuli: the idiosyncratic "high" and "low good doctor" faces, the "high" and "low good doctor" faces from the most dissimilar model among the other participants' models, and fixation cross. In each trial, the visual stimulus (idiosyncratic "high" / "low good doctor"; other's "high" / "low good doctor"; fixation cross) is presented for 5 seconds, followed by the delivery of an electrical stimulus. Participants then have 6.5 seconds to rate both the perceived pain intensity and unpleasantness (i.e., "How intense/unpleasant was the pain you felt?") on a 0 ("Not at all") to 100 ("Extremely") numeric scale.

Trial design is illustrated in Figure 4.2 (see Chapter 4, subchapter 4.2).

The task is structured in a block design with 5 blocks, each for condition: the idiosyncratic "high good doctor" block, the idiosyncratic "low good doctor" block, the other participant's "high good doctor" block; the other participant's "low good doctor" block; and the "fixation cross" block.

Each block consists of 6 trials lasting 20 seconds each, for a total block duration of 120 seconds. Block order is counterbalanced across participants. The total duration of the task is about 12 minutes.

iii) Familiarity Task. After completing the pain task, participants are administered PANAS for the second time and perform the familiarity rating task. They are asked to evaluate the 4 facial images displayed in the previous task and 4 new facial images (2 "high" and 2 "low good doctor" faces) selected from two other participants' models. Specifically, we select one very similar (CS closer to 1) and one dissimilar (CS closer to 0) model from theirs; resulting in 8 facial images. For each face, participants are instructed to report, on a numeric scale from 0 ("Not at all") to 100 ("Extremely"), how much it seems "familiar".

Two practice trials are administered to familiarize themselves with the procedure. Then, facial stimuli are presented one after the other.

Post-experimental questionnaire completion. Finally, participants are instructed to complete additional questionnaires within 24 hours of the experimental session, sent via email. The assessed personality traits include: i) trait anxiety (STAI-T; Spielberger, 1989, Italian adaptation by Pedrabissi & Santinello, 1989; possible range: 20-80, clinical cut off: 40), pain catastrophizing (Pain Catastrophizing Scale (PCS); Sullivan et al., 1995, Italian adaptation by Monticone et al., 2012; possible range: 0- 52), fear of pain (Fear of Pain Questionnaire; FPQ-III; McNeil & Rainwater, 1998, Italian adaptation by Tella et al., 2019; possible range: 30-150), the Big Five traits– extraversion, openness, agreeableness, conscientiousness, and emotional stability (BFI-10;Rammstedt & John, 2007; possible range: 2- 10 for each scale); and the behavioral inhibition/activation (BIS/BAS Scale; Carver & White, 1994, Italian adaptation by Leone et al., 2001; possible range BIS: 7- 28; Drive scale: 4-16; Fun: 4-16; Reward: 5-20). The influence of the physician's appearance on the participant was also assessed using the questionnaire described in Aitken et al., 2014. Importantly, we measured both the importance of the physician's physical appearance in terms of trust (item: "How important is the physician's appearance to you in terms of trust?" on a scale from 1 ("Not at all") to 5 ("Extremely") and the importance of physician's grooming (item: Please rank from 1 (more important) to 4 (least important) the following physician's characteristics: personal grooming, kindness, empathy, clinical knowledge. Participants who ranked personal grooming as the most important characteristic were coded as 1; all others were coded as 0).

All participants are subsequently fully debriefed regarding the study's aims, manipulated variables, hypotheses, and expected outcomes. If the participant was a student, they were awarded credits for their participation.

6.2.4 Preliminary data analysis

To assess test- retest reliability for each participant, we computed the correlation between the ratings participants gave to repeated stimuli. Participants with correlations less than or equal to 0.05 were excluded from subsequent analyses (Albohn et al., 2025).

6.2.4.1 Pain task

Pain measures: pain intensity and unpleasantness. Correlational analyses were conducted between mean pain intensity and unpleasantness in each condition and participants' pain expectations. To analyze data from the pain task, linear mixed- effects models (LMMs) were conducted separately for the two dependent variables, namely pain intensity and pain unpleasantness ratings.

The main effect of Condition (idiosyncratic "high good doctor"/ idiosyncratic "low good doctor"/other's "high good doctor"/other's "low good doctor"/ fixation cross) was tested specifying models including Condition as fixed effect and participant ID as a random intercept.

Fixed effects were evaluated using type-III ANOVAs on maximum-likelihood–fitted models. Model fit was evaluated using restricted maximum-likelihood–fitted models, and effect sizes were quantified with marginal and conditional R^2.

Post hoc comparisons of estimated marginal means (EMMs) were computed with Bonferroni correction for multiple testing.

Physiological measures: mean HR and HRV components. To test the effect of Condition on the mean HR and the parasympathetic and sympathetic components of HRV (i.e., RMSSD, VLF, LF, HF, HF/LF) calculated for each block, we followed the same procedure described in Study 2. We applied the following exclusion criteria on our sample: i) missing HR recording during the task performance (N=4); ii) Z scores for mean HR and HRV components <-3 and >3 (N=5). Therefore, analyses on physiological measures were conducted on a sample of 21 participants (11 women, 10 men, age range: 18-58; M=26.95; SD =8.73).

Heart rate was recorded as beat-to-beat intervals in ms with the Bodyguard 2 (Firstbeat Technologies Ltd.; Parak & Korhonen, 2013). HRV was derived by computing the root mean square of successive beat-to-beat interval differences (rMSSD), which reflects vagal regulation of HR with reduced susceptibility to respiratory influences (Laborde et al., 2017; Penttilä et al., 2001). Analyses were

performed using Kubios HRV software (Tarvainen et al., 2014), which uses an advanced detrending method based on the formulation of smoothing priorities (Tarvainen et al., 2002). Artifacts and ectopic beats were corrected using Kubios HRV software by two different methods: (1) A threshold-based correction, which is based on the comparison of each RR interval value to a local average interval; (2) an automatic correction, where artifacts are detected from a time series consisting of the differences between successive RR intervals.

Questionnaires. A paired samples *t*-test was performed on PANAS scores collected before and after the pain task to examine whether differences in reported positive and negative affective states emerged. Moreover, descriptive analyses were conducted on all questionnaires scoring to provide an overview of the sample characteristics. Finally, correlational analyses with the main dependent variables were also performed, to explore possible associations between personality traits and the experimental variables. Questionnaires scorings were not used as covariates in the main analyses due to the small sample size.

From the initial sample (N=30), only one participant was excluded from the analysis due to the lack of completion of the questionnaires. Therefore, these exploratory analyses were conducted on a sample of 29 participants (14 women, 15 men, range age:19-48 years, M=26.69, SD=7.64).

All analyses were performed in SPSS (version 25.0; Chicago, IL) and R (version 4.4.1; 2024-06-14 ucrt; Posit, PBC, Boston, MA, USA), using the lme4, emmeans and performance packages.

The alpha level for all analyses was set to p<.05.

6.2.4.2 Familiarity task

Data from the familiarity task were analyzed using a repeated measure ANOVA, with Condition (idiosyncratic “high”/ “low good doctor”; another participant’s “high” / “low good doctor”; similar model “high”/ “low good doctor”; dissimilar model “high”/ “low good doctor”) as a within-subject factor. Effect sizes were reported as partial eta squared (η^2_p). Post hoc pairwise comparisons were adjusted using Bonferroni correction.

Data preprocessing and stimulus generation were implemented in Python (v.3.10) using custom scripts and publicly available functions. Statistical analyses were instead performed in R (version 4.4.1; 2024-06-14 ucrt) using the lme4, emmeans and performance packages.

6.3 Preliminary results

Preliminary results are presented separately for the pain task and the familiarity task.

6.3.1 Pain task

6.3.1.1. Pain measures: pain intensity and unpleasantness

No significant correlations emerged between pain expectations and mean pain intensity and unpleasantness for each condition.

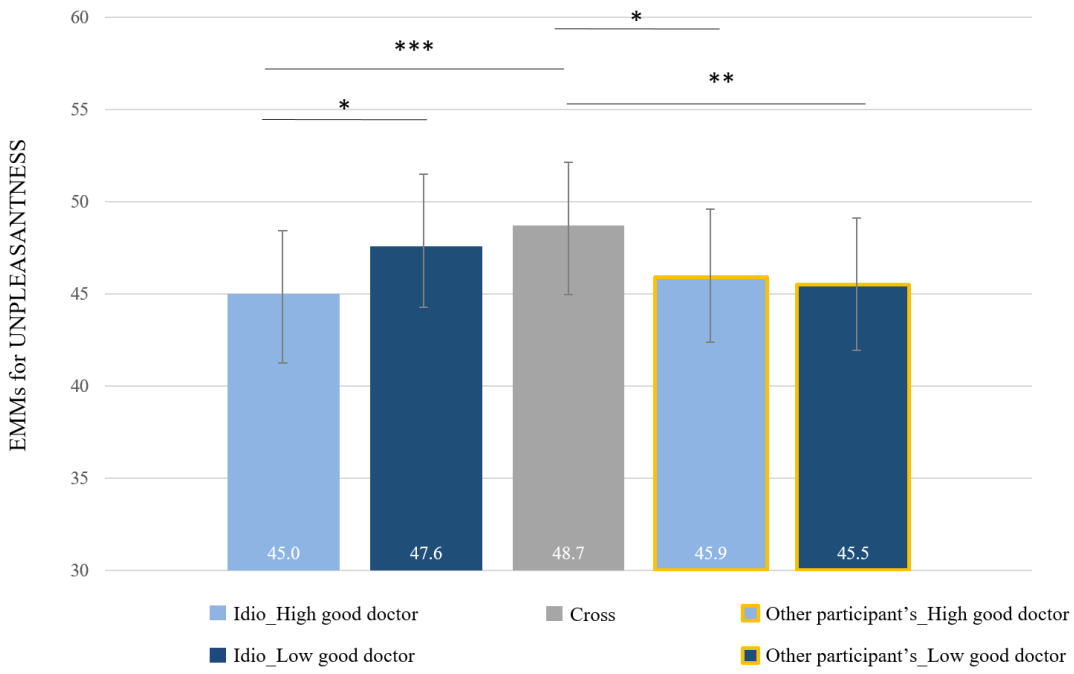
Linear mixed-effects models (LMMs) revealed that although no significant effects emerged for pain intensity ($F(4,866)=1.69$, $p=.15$, $R^2_c=.75$, $R^2_m=.002$), Condition (idiosyncratic "high"/ "low good doctor"; other participant's "high" / "low good doctor"; fixation cross) significantly affected participants' unpleasantness ($F(4,866)=6.05$, $p<.001$, $R^2_c=.82$, $R^2_m=.005$).

Specifically, estimated marginal means indicated that participants reported lower unpleasantness in the idiosyncratic "high good doctor" condition ($M=45.00$, 95% $CI[38.40, 51.70]$) compared to both the idiosyncratic "low good doctor" ($M=47.60$, 95% $CI[40.90, 54.30]$, $p<.05$) and fixation cross ($M=48.70$, 95% CI $[42.00, 55.40]$, $p<.001$). Moreover, participants also reported significantly lower unpleasantness when other participant's "high good doctor" ($M=45.90$, 95% $CI[39.30, 52.60]$, $p<.05$) and "low good doctor" faces ($M=45.50$, 95% $CI[38.80, 52.20]$, $p<.01$) were presented compared to the fixation cross.

Results are displayed in Figure 6.3

Figure 6.3

EMMs for pain unpleasantness across the idiosyncratic "high good doctor"/ "low good doctor", other participant's "high good doctor"/ "low good doctor" and fixation cross conditions.



Note. Error bars represent the standard error of the EMMs. *p<.05; **p<.01; ***p<.001

No other significant differences emerged. The complete output for the linear mixed-effect models and all the results are displayed in the Supplementary materials (see subchapter 8.3.1)

6.3.1.2 Physiological measures: Mean HR and HRV components

Repeated measure ANOVAs evidenced no significant main effect of Condition neither on Mean HR (F(2.66, 53.19)=1.36, p=.26, η^2_p=.06) nor on HRV components (RMSSD: F(1.89, 37.69)=.10, p=.89, η^2_p=.01; VLF: F(2.88, 57.64)=.22, p=.88, η^2_p=.01; LF: F(2.99, 59.92)=1.03, p=.39, η^2_p=.05; HF: F(2.05, 41.06)=.65, p=.59, η^2_p=.04; HF/LF: F(2.81, 56.28)=2.29, p=.09, η^2_p=.10).

Post-hoc comparison results are reported on the Supplementary materials.

6.3.2 Familiarity task

A repeated measure ANOVA evidenced a significant main effect of Condition (idiosyncratic "high good doctor"/ "low good doctor"; other participant's "high good doctor" / "low good doctor"; other

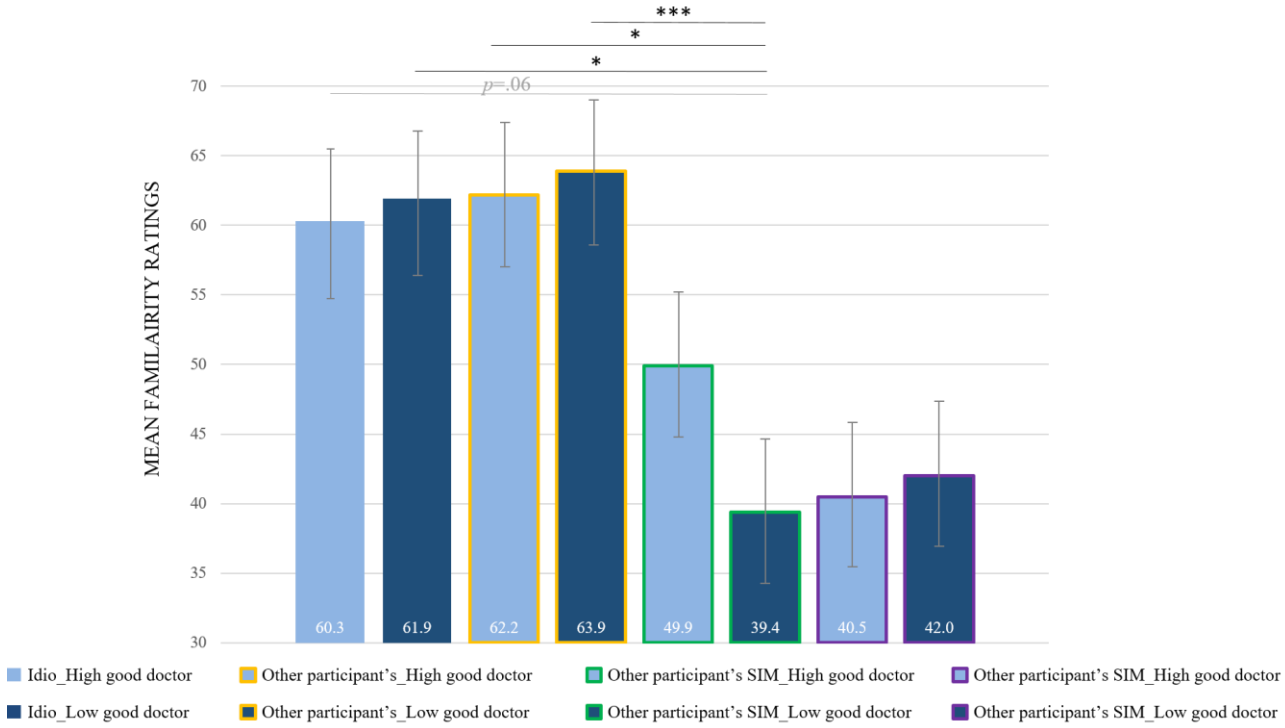
participant's "high good doctor"/ "low good doctor" similar model; other participant's "high good doctor"/ "low good doctor" dissimilar model) on participants' familiarity ratings ($F(4.23,122.60)=5.59$, $p<.001$, $\eta^2_p=.16$).

Analyses revealed that participants generally judged the stimuli they observed during the pain task as more familiar than the novel stimuli. Nonetheless, significant differences specifically emerged between "low good doctor" faces generated using the other participant's model considered as the most similar ($M=39.43$, $SD=24.42$) and all the facial stimuli observed during pain task ($M_{idio_high}=60.27$, $SD_{idio_high}=28.92$, $p=.06$; $M_{idio_low}=61.87$, $SD_{idio_low}=28.07$, $p<.05$; $M_{other_high}=62.20$, $SD_{other_high}=27.15$, $p<.05$; $M_{other_low}=63.90$, $SD_{other_low}=28.25$, $p<.01$).

Results are displayed in Figure 6.4.

Figure 6.4

Mean familiarity ratings across the different conditions, namely idiosyncratic "high good doctor"/ "low good doctor"; other participant's "high good doctor"/ "low good doctor"; other participant's similar "high good doctor"/ "low good doctor" and other participant's different "high good doctor"/ "low good doctor" conditions.



Note. Error bars represent the standard error of the means for each condition. *p<.05; **p<.01; ***p<.001

Non-significant results are reported in Supplementary materials (see subchapter 8.3.3, Table S8.16)

6.3.3 Questionnaires

A paired samples *t*-test on PANAS scoring collected before and after undergoing pain task showed that although no differences in positive affective states were observed ($t(28)=-0.62$, $p=.54$), participants reported significantly higher negative affect before ($M=15.76$;$SD=4.30$) than after ($M=12.34$, $SD=3.02$) the pain task ($t(28)=5.38$, $p<.001$).

Descriptive analyses on questionnaires scorings also evidenced that participants generally reported higher positive vs. negative affective states.

Results of descriptive analyses are displayed in Table 6.1.

Table 6.1

Descriptive statistics for the questionnaires administered during the whole experimental procedure.

	Minimum	Maximum	*M*	*SD*
PANAS				
Positive Affect pre pain	22	44	30.72	4.47
Negative Affect pre pain	10	28	15.76	4.30
Positive Affect post pain	20	41	31.10	5.41
Negative Affect post pain	10	19	12.34	3.02
STAI				
Trait Anxiety	27	58	45.07	8.12
State Anxiety	25	52	38.41	5.85
FPQ				
Fear of Pain	55	100	75.48	12.10
PCS				
Pain Catastrophizing	2	35	20.97	7.25
BFI				
Agreeableness	4	9	6.66	1.37
Conscientiousness	4	10	6.93	1.41
Emotional Stability	2	8	5.59	1.76
Extraversion	3	9	5.76	1.70
Openness	2	10	7.17	1.75
BIS/BAS				
Behavioural Inhibition	8	22	12.93	3.85
Drive	5	13	9.72	2.23
Fun	5	14	9.28	2.09
Reward	5	13	8.72	2.05
Importance trust	1	5	2.79	1.21

Correlational results evidenced significant positive correlations between mean pain intensity in the idiosyncratic “high good doctor” condition and Agreeableness trait (r=.38, p<.05), evidencing that the stronger the orientation toward interpersonal relationships, the higher the pain when participant’s

own “good doctor” is displayed. A positive correlation between other participant’s “high good doctor” condition and FPQ scoring (r=.39, p<.05) was also observed, suggesting that the greater the fear of pain, the stronger the perceived pain when the non-idiosyncratic “good doctor” representation is presented.

Significant negative correlations emerged, instead, between the mean pain intensity in the fixation cross, other participant’s “high” and “low good doctor” conditions and the Conscientiousness trait (r_{cross}=-.40, p<.05; r_{other_high}=-.37, p<.05;r_{other_low}=-.39, p<.05). Stated differently, more conscientious individuals report lower levels of pain when non-idiosyncratic “good doctor” faces, or no faces, are displayed. Lastly, negative correlations also emerged between mean pain intensity in the fixation cross and other participant’s “low good doctor” condition and Openness trait (r_{cross}=-.37, p=.05; r_{other_low}=-.37, p<.05; r_{other_low}=-.42, p<0.5). Hence, more open-minded individuals tend to report lower pain when faces reflecting different features from those they associate with their own “good doctor” stereotype, or no faces, are presented.

For pain unpleasantness, instead, a positive correlation emerged between the unpleasantness participants reported in the idiosyncratic “high good doctor” condition and the Drive scale (r=.37, p=.05); showing that more goal-oriented individuals tended to report higher unpleasantness when their own “good doctor” stereotype is displayed. Finally, consistent with previous results on pain intensity, a positive correlation also emerged between the mean unpleasantness in the other participant’s “high good doctor” condition and the Rumination scale (r=.39, p<.05), suggesting that higher levels of rumination may exacerbate the overall pain experience when non-idiosyncratic faces are presented.

Correlations between pain measures across all conditions and questionnaires are displayed in Table 6.2.

Table 6.2

Correlations between pain intensity and unpleasantness across all conditions and questionnaires

	Pain intensity				
	Idiosyncratic	**Idiosyncratic**	**Cross**	**Other**	**Other**
	High good doctor	**Low good doctor**		**High good doctor**	**Low good doctor**
PANAS					
Positive Affect pre pain	-.10	-.09	.30	-.14	-.32
Positive Affect post pain	.09	.10	.15	.10	.23
Negative Affect pre pain	-.05	-.10	-.31	-.01	-.17
Negative Affect post pain	.01	.01	.06	.06	.10
STAI					
Trait Anxiety	.15	.02	.26	.24	.31
State Anxiety	.08	-.04	.11	.20	.14
FPQ					
Fear of Pain	.21	.19	.22	.39*	.14
PCS					
Pain Catastorphising	.14	.08	.24	.25	.14
BFI					
Agreeableness	.38*	.26	.22	.31	.25
Conscientiousness	.-.35	-.30	-.40*	-.37*	-.39*
Emotional Stability	.09	.06	.24	.19	.19
Extraversion	-.04	.03	.02	-.07	-.05
Openness	-.29	-.21	-.37*	-.31	-.42*
BIS/BAS					

Behavioural inhibition	-.14	-.01	-.25	-.22	-.21
Drive	.28	.34	.26	.22	.25
Fun	.13	.16	.18	.11	.18
Reward	-.13	-.13	-.13	-.10	-.07
Importance trust	.01	.01	.001	-.13	-.12

	Unpleasantness				
	Idiosyncratic	**Idiosyncratic**	**Cross**	**Other**	**Other**
	High good doctor	**Low good doctor**		**High good doctor**	**Low good doctor**
PANAS					
Positive Affect pre pain	-.17	-.04	-.29	-.21	-.28
Positive Affect post pain	.09	.15	.08	.03	.11
Negative Affect pre pain	-.12	-.11	-.20	-.15	-.23
Negative Affect post pain	.03	.09	.07	.03	.06
STAI					
Trait Anxiety	.09	.02	.21	.16	.18
State Anxiety	.04	-.05	.05	.04	-.01
FPQ					
Fear of pain	.26	.18	.19	.35	.14
PCS					
Pain Catastrophizing	.19	.13	.21	.29	.13
BFI					
Agreeableness	.24	.14	.27	.17	.15
Conscientiousness	-.19	-.12	-.25	-.20	-.16

Emotional Stability	.13	.13	.12	.16	.14
Extraversion	-.03	.08	-.12	-.10	-.12
Openness	-.06	-.05	-.19	-.10	-.19
BIS/BAS					
Behavioural inhibition	-.16	-.12	-.21	-.24	-.25
Drive	.08	.03	.11	.12	.20
Fun	.37*	.27	.31	.35	.35
Reward	-.05	-.15	-.01	.01	.05
Importance trust	.01	.17	-.03	-.01	-.004

All correlations are displayed in the Supplementary materials (see subchapter 8.3.4, Table S8.17).

6.4 Discussion

In the present study, we provide additional evidence that complements and extends the findings reported in Study 2. Building on conclusions from Study 2- which demonstrated a modulatory effect of the stereotypical representation of a "good doctor" on the subjective experience of pain- and considering the relevance idiosyncratic contributions appear to have in shaping such stereotype at the individual level, we provide preliminary data supporting the hypothesis that faces aligning with the individual's own stereotypical mental representation of a "good doctor" modulate pain experience.

Data collection is still ongoing; therefore, the results reported here should be considered very preliminary and interpreted with caution, with emphasis placed on the discussion of outcomes related to the primary study objectives.

Specifically, our preliminary results indicate that, in line with the outcomes of Study 2, facial stimuli generated using each participant's idiosyncratic "good doctor" model effectively influenced the affective component of pain, i.e., unpleasantness, experienced during painful stimulations.

While acknowledging the limited sample size which precluded significant findings for pain intensity, we observed that idiosyncratic "high" and "low good doctor" faces correlated, respectively, with lower and higher unpleasantness levels. These results reinforce our previous findings, highlighting the robustness of such stereotypical representation in modulating the subjective experience of pain.

Notably, idiosyncratic “high good doctor” faces elicited the lowest unpleasantness, whereas the absence of facial stimuli– i.e., the fixation cross condition- was associated with the highest unpleasantness. Nonetheless, no differences emerged between the unpleasantness levels participants reported when idiosyncratic versus other participant’s “high” and “low good doctor” faces were presented. The data appear to suggest that the expected stronger pain modulation for idiosyncratic models is not clearly supported at this stage. However, given the limited sample size, these findings should be interpreted with caution, as they are preliminary and may not conclusively disconfirm the hypothesis.

No differences were found between the other participant’s “high” and “low good doctor” conditions either; suggesting that, irrespective of the type of face, participants simply benefited from the presentation of a facial stimulus compared to its absence.

Taken together, these findings indicate that although idiosyncratic “good doctor” models did not exert a significant stronger pain modulation compared to non-idiosyncratic models, they were still more effective in modulating the direction of perceived unpleasantness– reducing it when participant’s own stereotypical representation of a “good doctor” was administered and increasing it when faces characterized by features very distant from this representation were presented.

From a theoretical standpoint, these results suggest that idiosyncratic “high good doctor” faces may therefore act as particularly effective “safety signals” (Krahé et al., 2013), as they embody the social attributes- and the physical features conveying them- that each individual associates with their own “good doctor”. These faces may trigger positive feelings and a sense of security, which in turn may reduce pain (Coan et al., 2006; Krahé et al., 2013; Vlaeyen et al., 2009). Moreover, these findings align with the literature on pain modulation: indeed, unpleasantness has been considered as the affective-motivational component of pain- capturing the aspects that drives protective behaviours (Auvray et al., 2010). Moreover, previous findings show that emotions mainly influence pain unpleasantness, whereas conscious attention modulates pain intensity (Villemure & Bushnell, 2002, 2009). Finally, compared to pain intensity, unpleasantness is more strongly influenced by the social context in which the painful experience occurs (Fields, 1999; Leknes et al., 2013).

Lastly, results from the familiarity task evidenced that participants generally judged the stimuli already presented during the pain task as more familiar than those they never saw. However, faces manipulated using the model of another participant considered as the most similar were judged as the least familiar. These results are not surprising. Indeed, the selection of the most similar model among

those generated has been based on the cosine similarity computed between the vectors defining idiosyncratic models. Cosine similarity values closer to one indicate greater similarity between the two vectors in the latent space (Mana & Sasipraba, 2021). Specifically, values approaching one reflect a minimal angle between the vectors, meaning that they point in the same direction in the latent space. It follows that manipulations along these models produce highly comparable outcomes; and since "low good doctor" faces have been defined as the counterpart of the "high good doctor" faces- which are highly familiar- looking stimuli- its opposite would, by contrast, be perceived as unfamiliar.

This study has some limitations worth noting. First, sample size is smaller than that required to achieve adequate statistical power. Participant recruitment is still ongoing, and the findings reported here should be considered preliminary and interpreted with caution. Second, the facial stimuli administered to participants across the different tasks have not been balanced in terms of sex. Indeed, the only stimulus selection criterion has been the cosine similarity values calculated among the models. We are aware that face sex may represent an important aspect to be taken into account in social pain modulation in general and in pain modulation in clinical context in particular. Previous findings have shown that social contextual factors- as the sex of the experimenter– may differently affect participant's pain perception (Daniali & Flaten, 2019). Moreover, in the clinical setting it has been observed that patient-physician sex concordance is associated with better clinical outcomes (Heybati et al., 2025; Yao et al., 2022). Nevertheless, the primary aim of the present study is to determine whether idiosyncratic models of the "good doctor" stereotype would modulate pain perception more effectively than non-idiosyncratic representations. Since sex represents a fundamental component of each participant's internal representation of a "good doctor", we deliberately focused on testing the effects of this stereotype at the individual level rather than on the influence of different facial features on participant's painful experience. Future studies may further investigate whether manipulating the sex of facial stimuli within each participant's idiosyncratic model could differentially influence pain modulation.

6.5 Conclusions

In the present study, we showed that faces reflecting idiosyncratic representation of a "good doctor" – that is, faces embodying the specific facial features and personality traits each individual associates with this stereotype – modulates the subjective experience of pain.

The collected preliminary evidence contributes to reinforcing and extending previous findings on the social modulation of pain, especially within clinical context. Furthermore, these findings highlight the critical influence that social salient stimuli- as faces- exert on the subjective experience of pain.

Uncovering these mechanisms may ultimately pave the way for the development of more personalized and targeted interventions, fostering both theoretical advancements in pain research and improvements in patient care and recovery.

7. General conclusions

Pain has an intrinsic social nature (Sullivan & Williams, 2025), and its experience is shaped by the social context in which it occurs (Carlino et al., 2014; Hadjistavropoulos et al., 2011; Karos et al., 2018, 2020). Experiencing pain when conspecifics are present modulates pain perception, reflecting an evolutionary mechanism that facilitates coping when social support is perceived (Che et al., 2018; Krahé et al., 2013). Notably, perceived support can modulate the subjective experience of pain even when merely primed– as through the presentation of facial images of both affective others(Eisenberger et al., 2011; Hillmer et al., 2021) and strangers (Edwards et al., 2017).

Human faces are indeed powerful social stimuli conveying important information about conspecifics (Oosterhof & Todorov, 2008, 2009; Todorov et al., 2015, 2025). In the healthcare field, the implications of these facial- based social attributions have been observed to be particularly relevant, as they can affect patients' subjective experience of medical procedures, including satisfaction and pain intensity (Bagnis et al., 2023; Mattarozzi et al., 2021).

The present PhD project is positioned within this theoretical framework, bridging two major lines of investigation: the social modulation of pain and the impact of facial- appearance based inferences on individuals' behaviours. Leveraging on generative methods, we disentangled the physical and social traits people tend to associate with the stereotypical representation of a "good doctor"- that is, the stereotype a doctor should align with to be perceived as "good" at first glance- both at the group and the individual level. Then, we demonstrated how these representations modulate the subjective experience of pain, potentially reducing the sense of threat through primed safety and social support. Importantly, we particularly focused on the role that intrinsic familiarity inferred from facial appearance assumes in these processes, in light of evolutionary mechanisms that promote adaptive responses when conspecifics– especially those perceived as carrying safety signals – are present (Eisenberger et al., 2011; Krahé et al., 2013; Mattarozzi et al., 2021; Bagnis et al., 2023).

Specifically, the first study focused on identifying the physical features and social traits that people tend to associate with the stereotypical representation of a "good doctor." We developed a new dimension– namely, the "good doctor" dimension– along which facial stimuli could be manipulated to appear more or less similar to this stereotype. Consistent with recent literature (Hall et al., 2020; Necka et al., 2021), the "good doctor" dimension showed strong positive correlations with attributes related to four main social components: ethnicity and perceived similarity, attractiveness,

trustworthiness, and familiarity and professional look/intelligence. It also proved to be a valid tool for generating facial stimuli that reliably resemble this stereotype.

The second study, instead, investigated whether this stereotypical representation could attenuate the subjective experience of pain – reducing the sense of threat through inferences of familiarity. We further examined whether this modulatory effect may be influenced by individual differences, as genetic variations in polymorphisms involved in systems contributing to pain modulation and socio-affective behaviours and differences in HR and HRV. We demonstrated that faces aligning with the stereotypical representation of a "good doctor" are able to modulate pain perception. Nonetheless, this effect varies substantially across individuals according to their genetic profile, which could attenuate– or even reverse– this socially induced, placebo hypoalgesic effect. Specifically, individuals carrying allelic variants related to greater resilience to stress and lower anxiety proneness– namely, COMT GG (Serrano et al., 2019) and FAAH A* (Gärtner et al., 2019; Hariri et al., 2009), respectively – exhibited the strongest benefit from the presentation of faces reflecting the "good doctor" stereotype. Conversely, carriers of more dysregulated and less prosocial SNP variants, as TPH2 T* (Inoue et al., 2010) and OXTR-1 T* (LoParo & Waldman, 2015), reported lower pain levels when no facial stimuli were administered.

The third study explored more deeply how the "good doctor" stereotype is differently reflected in each individual. In agreement with the most recent literature on face evaluation (Albohn et al., 2022, 2024, 2025; Martinez et al., 2020), which highlights the non-negligible role of idiosyncratic over shared contributions in formulating complex social judgments, we developed idiosyncratic representations of the "good doctor". These representations appeared to reliably reflect this stereotype at the individual- level, showing stronger accuracy in predicting participants' judgments in terms of resemblance with a "good doctor" compared to non-idiosyncratic representations. Importantly, the idiosyncratic "good doctor" models also emerged to predict participants' judgments of perceived familiarity; corroborating the importance this attribute plays in shaping this stereotype. Furthermore, although sharing some commonalities, these models were very different from each other- confirming the relevance of idiosyncratic compared to shared contributions in formulating complex judgments (Albohn et al., 2024; Martinez et al., 2020).

Finally, the fourth study synthesized and consolidated the insights gathered throughout the project, investigating whether idiosyncratic "good doctor" representations – judged as more similar to a "good doctor" and, notably, more familiar –could modulate pain more effectively than non- idiosyncratic representations.

The results do not provide definitive evidence for the hypothesized greater pain modulation by idiosyncratic models; the limited sample size warrants cautious interpretation. Preliminary findings, based on a small sample, confirm a significant modulation of the pain affective component, such as unpleasantness, particularly at the idiosyncratic level (between high and low representations). However, these preliminary findings highlight the need for further investigation to better understand the phenomenon.

Together, despite some limitations – such as the deliberate decision to use a broad and unspecific label like "good doctor," intended to evoke the general stereotypical representation of a physician, or the relatively small sample sizes in the second and fourth studies – this project provides novel evidence for a form of social modulation of pain. Specifically, it demonstrates that facial stimuli imbued with positive valence, such as intrinsically familiar faces, can influence pain perception– an area that has been scarcely explored within healthcare contexts. In this framework, familiarity refers to the degree to which a face resembles someone known (Zebrowitz & Montepare, 2008; Zebrowitz et al., 2007). This intrinsic familiarity functions as a "safety signal," indicating environmental security and reducing threat vigilance, particularly in socially relevant situations such as pain induction (Krahé et al., 2013; Vlaeyen et al., 2009). In a clinical context, social stimuli like familiar faces may indeed evoke feelings of security and reassurance, leading to pain reduction. This may imply that, while pain functions as an essential alarm for protective actions, its intensity could be mitigated in environments that convey safety signals. Extending this line of inquiry, our investigation into individual differences in the mental representations of the "good doctor" occupational stereotype, including individual differences in specific SNPs and physiological measures such as HR and HRV, reveals that these mechanisms contribute to shaping individuals' pain experience, potentially providing an evolutionary advantage for some individuals but not others.

Additionally, this doctoral project holds the unique advantage of capitalizing on computational, data-driven methods. We have developed a model that not only visualizes such stereotypical representation, but also offers a valid tool to investigate its influence on clinically relevant outcomes. Furthermore, through the implementation of generative models, we have highlighted the relevance of idiosyncratic contributions in shaping occupational stereotypes, validating their usage in exploring more deeply how individuals differ in representing social roles through their own mind's eye. Together, this work contributes to expanding current knowledge on inferential processes in medical contexts and opens new avenues for integrating social-cognitive and computational approaches in healthcare research.

The translational potential of this research opens promising avenues for application within healthcare contexts. Understanding the social traits that patients predominantly associate with a “good doctor” – and how these traits influence both pain perception and treatment satisfaction– could inform new strategies to enhance clinician–patient communication and optimize the social environment of medical care. Such insights may, for instance, help increase clinicians’ awareness of patients’ expectations and of how these expectations shape their interpersonal behaviour during medical encounters. Moreover, leveraging familiarity-based mechanisms could inspire innovative interventions aimed at reducing pain and anxiety by enhancing perceived social safety– for example, through avatars or virtual agents that resemble familiar features or embody the “good doctor” stereotype, integrated into clinical or virtual care settings.

Altogether, these directions emphasize the importance of integrating social-cognitive, affective, and computational perspectives to advance our understanding of how social context shapes pain experience. By extending the theoretical, methodological, and translational implications of this work, future studies may contribute to the development of more personalized and socially attuned approaches to pain management in healthcare settings.

References

Adams, K., Hean, S., Sturgis, P., & Clark, J. M. (2006). Investigating the factors influencing professional identity of first-year health and social care students. *Learning in Health and Social Care*, *5*(2), 55–68. https://doi.org/10.1111/j.1473-6861.2006.00119.x

Aitken, S. A., Tinning, C. G., Gupta, S., Medlock, G., Wood, A. M., & Aitken, M. A. (2014). The importance of the orthopaedic doctors' appearance: A cross-regional questionnaire based study. *The Surgeon*, *12*(1), 40–46. https://doi.org/10.1016/j.surge.2013.07.002

Albohn, D. N., Martinez, J. E., & Todorov, A. (2024). Determinants of shared and idiosyncratic contributions to judgments of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *50*(11), 1117–1130. https://doi.org/10.1037/xhp0001239

Albohn, D. N., Uddenberg, S., & Todorov, A. (2022). A data-driven, hyper-realistic method for visualizing individual mental representations of faces. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.997498

Albohn, D. N., Uddenberg, S., & Todorov, A. (2025). Individualized models of social judgments and context-dependent representations. *Scientific Reports*, *15*(1), 4208. https://doi.org/10.1038/s41598-025-86056-1

Anderson, C., & Kilduff, G. J. (2009). Why do dominant personalities attain influence in face-to-face groups? The competence-signaling effects of trait dominance. *Journal of Personality and Social Psychology*, *96*(2), 491–503. https://doi.org/10.1037/a0014201

Arce, E., Simmons, A. N., Stein, M. B., Winkielman, P., Hitchcock, C., & Paulus, M. P. (2009). Association between individual differences in self-reported emotional resilience and the affective perception of neutral faces. *Journal of Affective Disorders*, *114*(1), 286–293. https://doi.org/10.1016/j.jad.2008.08.015

Ashton, M. C., & Lee, K. (2021). A review of personality/religiousness associations. *Current Opinion in Psychology*, *40*, 51–55. https://doi.org/10.1016/j.copsyc.2020.08.023

Ashton-James, C. E., Tybur, J. M., Grießer, V., & Costa, D. (2019). Stereotypes about surgeon warmth and competence: The role of surgeon gender. *PLOS ONE*, *14*(2), e0211890. https://doi.org/10.1371/journal.pone.0211890

Atlas, L. Y. (2021). A social affective neuroscience lens on placebo analgesia. *Trends in Cognitive Sciences*, *25*(11), 992–1005. https://doi.org/10.1016/j.tics.2021.07.016

Atlas, L. Y., & Wager, T. D. (2012). How expectations shape pain. *Neuroscience Letters*, *520*(2), 140–148. https://doi.org/10.1016/j.neulet.2012.03.039

Au, S., Khandwala, F., & Stelfox, H. T. (2013). Physician Attire in the Intensive Care Unit and Patient Family Perceptions of Physician Professional Characteristics. *JAMA Internal Medicine*, *173*(6), 465–467. https://doi.org/10.1001/jamainternmed.2013.2732

Auvray, M., Myin, E., & Spence, C. (2010). The sensory-discriminative and affective-motivational aspects of pain. *Neuroscience & Biobehavioral Reviews*, *34*(2), 214–223. https://doi.org/10.1016/j.neubiorev.2008.07.008

Bagnis, A., Caffo, E., Cipolli, C., De Palma, A., Farina, G., & Mattarozzi, K. (2020). Judging health care priority in emergency situations: Patient facial appearance matters. *Social Science & Medicine*, *260*, 113180.

Bagnis, A., Celeghin, A., Diano, M., Mendez, C. A., Spadaro, G., Mosso, C. O., Avenanti, A., & Tamietto, M. (2020). Functional neuroanatomy of racial categorization from visual perception: A meta-analytic study. *NeuroImage*, *217*, 116939. https://doi.org/10.1016/j.neuroimage.2020.116939

Bagnis, A., Celeghin, A., Mosso, C. O., & Tamietto, M. (2019). Toward an integrative science of social vision in intergroup bias. *Neuroscience and Biobehavioral Reviews*, *102*, 318–326. https://doi.org/10.1016/j.neubiorev.2019.04.020

Bagnis, A., Colonnello, V., Russo, P. M., & Mattarozzi, K. (2024). Facial trustworthiness dampens own-gender bias in emotion recognition. *Psychological Research*, *88*(2), 458–465. https://doi.org/10.1007/s00426-023-01864-2

Bagnis, A., Cremonini, V., Pasi, E., Pasquinelli, G., Rubbi, I., Russo, P. M., & Mattarozzi, K. (2021). Facing up to bias in healthcare: The influence of familiarity appearance on hiring decisions. *Applied Cognitive Psychology*, *35*(6), 1585–1591. https://doi.org/10.1002/acp.3873

Bagnis, A., Meeuwis, S. H., Haas, J. W., O'Keeffe, M., Bajcar, E. A., Babel, P., Evers, A. W. M., Glogan, E., Oleszczyk, M., Portoles, A., Vlaeyen, J. W. S., Mattarozzi, K., & PANACEA Consortium. (2025). A scoping review of placebo and nocebo responses and effects: Insights for clinical trials and practice. *Health Psychology Review*, *19*(2), 409–447. https://doi.org/10.1080/17437199.2025.2471792

Bagnis, A., Todorov, A., Altizio, I., Colonnello, V., Fanti, S., Russo, P. M., & Mattarozzi, K. (2023). Familiarity From Facial Appearance Leads to Hypoalgesia. *The Journal of Pain*, *24*(11), 2040–2051. https://doi.org/10.1016/j.jpain.2023.06.012

Balconi, M., Brambilla, E., & Falbo, L. (2009). Appetitive vs. Defensive responses to emotional cues. Autonomic measures and brain oscillation modulation. *Brain Research*, *1296*, 72–84. https://doi.org/10.1016/j.brainres.2009.08.056

Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences*, *104*(46), 17948–17953. https://doi.org/10.1073/pnas.0705435104

Barger, L. K., Ayas, N. T., Cade, B. E., Cronin, J. W., Rosner, B., Speizer, F. E., & Czeisler, C. A. (2006). Impact of Extended-Duration Shifts on Medical Errors, Adverse Events, and Attentional Failures. *PLOS Medicine*, *3*(12), e487. https://doi.org/10.1371/journal.pmed.0030487

Becker, E. R., & Roblin, D. W. (2008). Translating Primary Care Practice Climate into Patient Activation: The Role of Patient Trust in Physician. *Medical Care*, *46*(8), 795. https://doi.org/10.1097/MLR.0b013e31817919c0

Behrmann, M., & Avidan, G. (2022). Face perception: Computational insights from phylogeny. *Trends in Cognitive Sciences*, *26*(4), 350–363. https://doi.org/10.1016/j.tics.2022.01.006

Benedetti, F., Amanzio, M., Rosato, R., & Blanchard, C. (2011). Nonopioid placebo analgesia is mediated by CB1 cannabinoid receptors. *Nature Medicine*, *17*(10), 1228–1230. https://doi.org/10.1038/nm.2435

Bin Meshar, M., Stolier, R. M., & Freeman, J. B. (2022). *Facial Stereotyping Drives Judgments of Perceptually Ambiguous Social Groups* -.

https://journals.sagepub.com/doi/10.1177/1948550611062285?icid=int.sj-full-text.citing-articles.17

Birkhäuer, J., Gaab, J., Kossowsky, J., Hasler, S., Krummenacher, P., Werner, C., & Gerger, H. (2017). Trust in the health care professional and health outcome: A meta-analysis. *PLOS ONE*, *12*(2), e0170988. https://doi.org/10.1371/journal.pone.0170988

Bjornsdottir, R. T., Hehman, E., & Human, L. J. (2022). Consensus Enables Accurate Social Judgments. *Social Psychological and Personality Science*, *13*(6), 1010–1021. https://doi.org/10.1177/19485506211047095

Borracci, R. A., Gallesio, J. M. Á., Ciambrone, G., Matayoshi, C., Rossi, F., & Cabrera, S. (2020). What patients consider to be a 'good' doctor, and what doctors consider to be a 'good' patient: A text-mining algorithm-based analysis. *Revista Médica de Chile*, *148*(7), Article 7. https://www.revistamedicadechile.cl/index.php/rmedica/article/view/8307

Brączyk, J., & Bąbel, P. (2021). The Role of the Observers' Perception of a Model's Self-Confidence in Observationally Induced Placebo Analgesia. *The Journal of Pain*, *22*(12), 1672–1680. https://doi.org/10.1016/j.jpain.2021.06.001

Bridges, D., Wulff, E., Bamberry, L., Krivokapic-Skoko, B., & Jenkins, S. (2020). Negotiating gender in the male-dominated skilled trades: A systematic literature review. *Construction Management and Economics*, *38*(10), 894–916. https://doi.org/10.1080/01446193.2020.1762906

Brown, J. L., Sheffield, D., Leary, M. R., & Robinson, M. E. (2003). Social support and experimental pain. *Psychosomatic Medicine*, *65*(2), 276–283. https://doi.org/10.1097/01.psy.0000030388.62434.46

Brown, S. M., Peet, E., Manuck, S. B., Williamson, D. E., Dahl, R. E., Ferrell, R. E., & Hariri, A. R. (2005). A regulatory variant of the human tryptophan hydroxylase-2 gene biases amygdala reactivity. *Molecular Psychiatry*, *10*(9), 884–888, 805. https://doi.org/10.1038/sj.mp.4001716

Brüne, M. (2012). Does the oxytocin receptor polymorphism (rs2254298) confer "vulnerability" for psychopathology or "differential susceptibility"? Insights from evolution. *BMC Medicine*, *10*(1), 38. https://doi.org/10.1186/1741-7015-10-38

Burford, B., & Rosenthal-Stott, H. E. S. (2017). First and second year medical students identify and self-stereotype more as doctors than as students: A questionnaire study. *BMC Medical Education*, *17*(1), 209. https://doi.org/10.1186/s12909-017-1049-2

Burgess, C., Klemt Craig, C. B., & Klemmer, C. L. (2021). Sexual and Gender Minority Marginalization in Military Contexts. In E. M. Lund, C. Burgess, & A. J. Johnson (Eds.), *Violence Against LGBTQ+ Persons: Research, Practice, and Advocacy* (pp. 309–316). Springer International Publishing. https://doi.org/10.1007/978-3-030-52612-2_24

Butkus, R., Serchen, J., Moyer, D. V., Bornstein, S. S., Hingle, S. T., & for the Health and Public Policy Committee of the American College of Physicians. (2018). Achieving Gender Equity in Physician Compensation and Career Advancement: A Position Paper of the American College of Physicians. *Annals of Internal Medicine*, *168*(10), 721–723. https://doi.org/10.7326/M17-3438

Cajanus, K., Holmström, E. J., Wessman, M., Anttila, V., Kaunisto, M. A., & Kalso, E. (2016). Effect of endocannabinoid degradation on pain: Role of: FAAH: polymorphisms in experimental and postoperative pain in women treated for breast cancer. *PAIN*, *157*(2), 361. https://doi.org/10.1097/j.pain.0000000000000398

Carlino, E., Frisaldi, E., & Benedetti, F. (2014). Pain and the context. *Nature Reviews. Rheumatology*, *10*(6), 348–355. https://doi.org/10.1038/nrrheum.2014.17

Carr, E. W., Brady, T. F., & Winkielman, P. (2017). Are You Smiling, or Have I Seen You Before? Familiarity Makes Faces Look Happier. *Psychological Science*, *28*(8), 1087–1102. https://doi.org/10.1177/0956797617702003

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Castelli, L., Carraro, L., Ghitti, C., & Pastore, M. (2009). The effects of perceived competence and sociability on electoral outcomes. *Journal of Experimental Social Psychology*, *45*(5), 1152–1155. https://doi.org/10.1016/j.jesp.2009.06.018

Ceccarelli, I., Bagnis, A., Ottaviani, C., Thayer, J. F., & Mattarozzi, K. (2024). Racial biases, facial trustworthiness, and resting heart rate variability: Unravelling complexities in pain recognition.

Cognitive Research: Principles and Implications, *9*(1), 69. https://doi.org/10.1186/s41235-024-00588-0

Chang, L. J., Doll, B. B., van 't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, *61*(2), 87–105. https://doi.org/10.1016/j.cogpsych.2010.03.001

Che, X., Cash, R., Chung, S., Fitzgerald, P. B., & Fitzgibbon, B. M. (2018). Investigating the influence of social support on experimental pain and related physiological arousal: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *92*, 437–452. https://doi.org/10.1016/j.neubiorev.2018.07.005

Che, X., Cash, R., Fitzgerald, P., & Fitzgibbon, B. M. (2018). The Social Regulation of Pain: Autonomic and Neurophysiological Changes Associated With Perceived Threat. *The Journal of Pain*, *19*(5), 496–505. https://doi.org/10.1016/j.jpain.2017.12.007

Chemers, M. (2014). *An Integrative Theory of Leadership*. Psychology Press. https://doi.org/10.4324/9781315805726

Chen, G.-L., Vallender, E. J., & Miller, G. M. (2008). Functional characterization of the human TPH2 5′ regulatory region: Untranslated region and polymorphisms modulate gene expression in vitro. *Human Genetics*, *122*(6), 645–657. https://doi.org/10.1007/s00439-007-0443-y

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*(1), 103–125. https://doi.org/10.1037/a0030398

Chiang, K. P., Gerber, A. L., Sipe, J. C., & Cravatt, B. F. (2004). Reduced cellular expression and activity of the P129T mutant of human fatty acid amide hydrolase: Evidence for a link between defects in the endocannabinoid system and problem drug use. *Human Molecular Genetics*, *13*(18), 2113–2119. https://doi.org/10.1093/hmg/ddh216

Chung, H., Lee, H., Chang, D.-S., Kim, H.-S., Lee, H., Park, H.-J., & Chae, Y. (2012). Doctor's attire influences perceived empathy in the patient–doctor relationship. *Patient Education and Counseling*, *89*(3), 387–391. https://doi.org/10.1016/j.pec.2012.02.017

Coan, J. A., & Sbarra, D. A. (2015). *Social Baseline Theory*: The social regulation of risk and effort. *Current Opinion in Psychology*, *1*, 87–91. https://doi.org/10.1016/j.copsyc.2014.12.021

Coan, J. A., Schaefer, H. S., & Davidson, R. J. (2006). Lending a hand: Social regulation of the neural response to threat. *Psychological Science*, *17*(12), 1032–1039. https://doi.org/10.1111/j.1467-9280.2006.01832.x

Colagiuri, B., Schenk, L. A., Kessler, M. D., Dorsey, S. G., & Colloca, L. (2015). The placebo effect: From concepts to genes. *Neuroscience*, *307*, 171–190. https://doi.org/10.1016/j.neuroscience.2015.08.017

Colloca, L., & Barsky, A. J. (2020). Placebo and Nocebo Effects. *The New England Journal of Medicine*, *382*(6), 554–561. https://doi.org/10.1056/NEJMra1907805

Colloca, L., Panaccione, R., & Murphy, T. K. (2019). The Clinical Implications of Nocebo Effects for Biosimilar Therapy. *Frontiers in Pharmacology*, *10*. https://doi.org/10.3389/fphar.2019.01372

Colloca, L., Wang, Y., Martinez, P. E., Chang, Y.-P. C., Ryan, K. A., Hodgkinson, C., Goldman, D., & Dorsey, S. G. (2019). OPRM1 rs1799971, COMT rs4680, and FAAH rs324420 genes interact with placebo procedures to induce hypoalgesia. *PAIN*, *160*(8), 1824. https://doi.org/10.1097/j.pain.0000000000001578

Colonnello, V., Russo, P. M., & Mattarozzi, K. (2019). First Impression Misleads Emotion Recognition. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00527

Cook, K., & Stoecker, J. (2014). Healthcare Student Stereotypes: A Systematic Review with Implications for Interprofessional Collaboration. *Journal of Research in Interprofessional Practice and Education*, *4*(2). https://doi.org/10.22230/jripe.2014v4n2a151

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews. Neuroscience*, *3*(8), 655–666. https://doi.org/10.1038/nrn894

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map. In *Advances in Experimental Social Psychology* (Vol. 40, pp. 61–149). Academic Press. https://doi.org/10.1016/S0065-2601(07)00002-0

Cuddy, F., Susan T. (2004). *Ageism: Stereotyping and Prejudice Against Older Persons*. MIT Press.

Daniali, H., & Flaten, M. A. (2019). A Qualitative Systematic Review of Effects of Provider Characteristics and Nonverbal Behavior on Pain, and Placebo and Nocebo Effects. *Frontiers in Psychiatry*, *10*. https://doi.org/10.3389/fpsyt.2019.00242

Day, M. A., Ward, L. C., de la Vega, R., Ehde, D. M., & Jensen, M. P. (2022). Development of the Pain Responses Scale: A measure informed by the BIS-BAS model of pain. *European Journal of Pain*, *26*(2), 505–521. https://doi.org/10.1002/ejp.1877

de C Williams, A. C. (2023). Pain. *Evolution, Medicine, and Public Health*, *11*(1), 429–437. https://doi.org/10.1093/emph/eoad038

de C Williams, A. C., & Craig, K. D. (2006). A science of pain expression? *Pain*, *125*(3), 202–203. https://doi.org/10.1016/j.pain.2006.08.004

de C Williams, A. C., Gallagher, E., Fidalgo, A. R., & Bentley, P. J. (2016). Pain expressiveness and altruistic behavior: An exploration using agent-based modeling. *PAIN*, *157*(3), 759. https://doi.org/10.1097/j.pain.0000000000000443

Debiec, J., & Sullivan, R. M. (2017). The neurobiology of safety and threat learning in infancy. *Neurobiology of Learning and Memory*, *143*, 49–58. https://doi.org/10.1016/j.nlm.2016.10.015

DeBruine, L. M. (2002). Facial resemblance enhances trust. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1498), 1307–1312. https://doi.org/10.1098/rspb.2002.2034

Dincheva, I., Drysdale, A. T., Hartley, C. A., Johnson, D. C., Jing, D., King, E. C., Ra, S., Gray, J. M., Yang, R., DeGruccio, A. M., Huang, C., Cravatt, B. F., Glatt, C. E., Hill, M. N., Casey, B. J., & Lee, F. S. (2015). FAAH genetic variation enhances fronto-amygdala function in mouse and human. *Nature Communications*, *6*(1), 6395. https://doi.org/10.1038/ncomms7395

Diniz, D. A., Petrocchi, J. A., Navarro, L. C., Souza, T. C., Castor, M. G. M. e, Duarte, I. D. G., & Romero, T. R. L. (2018). Serotonin induces peripheral antinociception via the opioidergic system. *Biomedicine & Pharmacotherapy*, *97*, 1434–1437. https://doi.org/10.1016/j.biopha.2017.11.048

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*(3), 285–290. https://doi.org/10.1037/h0033731

Dion, K. K. (2002). Cultural perspectives on facial attractiveness. In *Facial attractiveness: Evolutionary, cognitive, and social perspectives* (pp. 239–259). Ablex Publishing.

Doi, H., Nishitani, S., & Shinohara, K. (2015). Association between catechol-*O*-methyltransferase Val158Met polymorphism and configural mode of face processing. *Neuroscience Letters, 586*, 19–23. https://doi.org/10.1016/j.neulet.2014.12.001

Dotsch, R., Hassin, R. R., & Todorov, A. (2016). Statistical learning shapes face evaluation. *Nature Human Behaviour, 1*(1), 1–6. https://doi.org/10.1038/s41562-016-0001

Drabant, E. M., Hariri, A. R., Meyer-Lindenberg, A., Munoz, K. E., Mattay, V. S., Kolachana, B. S., Egan, M. F., & Weinberger, D. R. (2006). Catechol O-methyltransferase Val158Met Genotype and Neural Mechanisms Related to Affective Arousal and Regulation. *Archives of General Psychiatry, 63*(12), 1396–1406. https://doi.org/10.1001/archpsyc.63.12.1396

Dumas, R., & Testé, B. (2006). The Influence of Criminal Facial Stereotypes on Juridic Judgments. *Swiss Journal of Psychology*, *65*(4), 237–244. https://doi.org/10.1024/1421-0185.65.4.237

Ebert, A., & Brüne, M. (2018). Oxytocin and Social Cognition. In R. Hurlemann & V. Grinevich (Eds.), *Behavioral Pharmacology of Neuropeptides: Oxytocin* (pp. 375–388). Springer International Publishing. https://doi.org/10.1007/7854_2017_21

Edwards, R., Eccleston, C., & Keogh, E. (2017). Observer influences on pain: An experimental series examining same-sex and opposite-sex friends, strangers, and romantic partners. *PAIN*, *158*(5), 846. https://doi.org/10.1097/j.pain.0000000000000840

Eisenberger, N. I., Master, S. L., Inagaki, T. K., Taylor, S. E., Shirinyan, D., Lieberman, M. D., & Naliboff, B. D. (2011). Attachment figures activate a safety signal-related neural region and reduce pain experience. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(28), 11721–11726. https://doi.org/10.1073/pnas.1108239108

Emerick, T., Alter, B., Jarquin, S., Brancolini, S., Bernstein, C., Luong, K., Morrissey, S., & Wasan, A. (2020). Telemedicine for Chronic Pain in the COVID-19 Era and Beyond. *Pain Medicine*, *21*(9), 1743–1748. https://doi.org/10.1093/pm/pnaa220

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Ferguson, M. J., & Bargh, J. A. (2004). How social perception can automatically influence behavior. *Trends in Cognitive Sciences*, *8*(1), 33–39. https://doi.org/10.1016/j.tics.2003.11.004

Fields, H. L. (1999). Pain: An unpleasant topic. *Pain*, *82*, S61–S69. https://doi.org/10.1016/S0304-3959(99)00139-6

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Friedman, S., & Laurison, D. (2020). The Class Ceiling: Why it Pays to be Privileged. *Social Forces*, *99*(1), e5. https://doi.org/10.1093/sf/soz170

Friedmann, E., & Efrat-Treister, D. (2023). Gender Bias in Stem Hiring: Implicit In-Group Gender Favoritism Among Men Managers. *Gender & Society*, *37*(1), 32–64. https://doi.org/10.1177/08912432221137910

Frot, M., Feine, J. S., & Bushnell, C. M. (2004). Sex differences in pain perception and anxiety. A psychophysical study with topical capsaicin. *Pain*, *108*(3), 230–236. https://doi.org/10.1016/j.pain.2003.11.017

Fujita, K., Fishbach, A., & Liberman, N. (2025). *Psychological Quest for Meaning*. Guilford Publications.

Fullana, M. A., Harrison, B. J., Soriano-Mas, C., Vervliet, B., Cardoner, N., Àvila-Parcet, A., & Radua, J. (2016). Neural signatures of human fear conditioning: An updated and extended meta-analysis of fMRI studies. *Molecular Psychiatry*, *21*(4), 500–508. https://doi.org/10.1038/mp.2015.88

Furmark, T., Appel, L., Henningsson, S., Åhs, F., Faria, V., Linnman, C., Pissiota, A., Frans, Ö., Bani, M., Bettica, P., Pich, E. M., Jacobsson, E., Wahlstedt, K., Oreland, L., Långström, B., Eriksson, E., & Fredrikson, M. (2008). A Link between Serotonin-Related Gene Polymorphisms, Amygdala Activity, and Placebo-Induced Relief from Social Anxiety. *The Journal of Neuroscience*, *28*(49), 13066–13074. https://doi.org/10.1523/JNEUROSCI.2534-08.2008

Gao, J., Pan, Z., Jiao, Z., Li, F., Zhao, G., Wei, Q., Pan, F., & Evangelou, E. (2012). TPH2 Gene Polymorphisms and Major Depression – A Meta-Analysis. *PLOS ONE*, *7*(5), e36721. https://doi.org/10.1371/journal.pone.0036721

Gärtner, A., Dörfel, D., Diers, K., Witt, S. H., Strobel, A., & Brocke, B. (2019). Impact of FAAH genetic variation on fronto-amygdala function during emotional processing. *European Archives of Psychiatry and Clinical Neuroscience*, *269*(2), 209–221. https://doi.org/10.1007/s00406-018-0944-9

Goldman, D., Oroszi, G., & Ducci, F. (2006). The Genetics of Addictions: Uncovering the Genes. *Focus*, *4*(3), 401–415. https://doi.org/10.1176/foc.4.3.401

Gondo, M., Moriguchi, Y., Kodama, N., Sato, N., Sudo, N., Kubo, C., & Komaki, G. (2012). Daily physical complaints and hippocampal function: An fMRI study of pain modulation by anxiety. *NeuroImage*, *63*(3), 1011–1019. https://doi.org/10.1016/j.neuroimage.2012.07.025

Gordon, R. A., & Arvey, R. D. (2004). Age Bias in Laboratory and Field Settings: A Meta-Analytic Investigation. *Journal of Applied Social Psychology*, *34*(3), 468–492. https://doi.org/10.1111/j.1559-1816.2004.tb02557.x

Graham, J. R., Harvey, C. R., & Puri, M. (2017). A corporate beauty contest. *Management Science*, *63*(9), 3044–3056. https://doi.org/10.1287/mnsc.2016.2484

Grün, F. C., Heibges, M., Westfal, V., & Feufel, M. A. (2022). "You Never Get a Second Chance": First Impressions of Physicians Depend on Their Body Posture and Gender. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.836157

Guo, Y., Logan, H. L., Glueck, D. H., & Muller, K. E. (2013). Selecting a sample size for studies with repeated measures. *BMC Medical Research Methodology*, *13*, 100. https://doi.org/10.1186/1471-2288-13-100

Hadjistavropoulos, T., Craig, K. D., Duck, S., Cano, A., Goubert, L., Jackson, P. L., Mogil, J. S., Rainville, P., Sullivan, M. J. L., Williams, A. C. de C., Vervoort, T., & Fitzgerald, T. D. (2011). A biopsychosocial formulation of pain communication. *Psychological Bulletin*, *137*(6), 910–939. https://doi.org/10.1037/a0023876

Hall, J. A., Ruben, M. A., & Swatantra. (2020). First Impressions of Physicians According to Their Physical and Social Group Characteristics. *Journal of Nonverbal Behavior*, *44*(2), 279–299. https://doi.org/10.1007/s10919-019-00329-8

Hall, K. T., Lembo, A. J., Kirsch, I., Ziogas, D. C., Douaiher, J., Jensen, K. B., Conboy, L. A., Kelley, J. M., Kokkotou, E., & Kaptchuk, T. J. (2012). Catechol-O-methyltransferase val158met polymorphism predicts placebo effect in irritable bowel syndrome. *PloS One*, *7*(10), e48135. https://doi.org/10.1371/journal.pone.0048135

Hall, K. T., Loscalzo, J., & Kaptchuk, T. J. (2015). Genetics and the placebo effect: The placebome. *Trends in Molecular Medicine*, *21*(5), 285–294. https://doi.org/10.1016/j.molmed.2015.02.009

Han, J., Kamber, M., & Pei, J. (2012). 2—Getting to Know Your Data. In J. Han, M. Kamber, & J. Pei (Eds.), *Data Mining: Concepts and Techniques (Third Edition)* (pp. 39–82). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-381479-1.00002-2

Hareli, S., & Hess, U. (2012). The social signal value of emotions. *Cognition & Emotion*, *26*(3), 385–389. https://doi.org/10.1080/02699931.2012.665029

Harendza, S., & Pyra, M. (2017). Just fun or a prejudice? – Physician stereotypes in common jokes and their attribution to medical specialties by undergraduate medical students. *BMC Medical Education*, *17*(1), 128. https://doi.org/10.1186/s12909-017-0964-6

Hariri, A. R., Gorka, A., Hyde, L. W., Kimak, M., Halder, I., Ducci, F., Ferrell, R. E., Goldman, D., & Manuck, S. B. (2009). Divergent Effects of Genetic Variation in Endocannabinoid Signaling on Human Threat- and Reward-Related Brain Function. *Biological Psychiatry*, *66*(1), 9–16. https://doi.org/10.1016/j.biopsych.2008.10.047

Hatt, B. (2012). Smartness as a Cultural Practice in Schools. *American Educational Research Journal*, *49*(3), 438–460. https://doi.org/10.3102/0002831211415661

Haynes, D., Pampari, A., Topham, C., Schwarzenberger, K., Heath, M., Zou, J., & Greiling, T. M. (2021). Patient Experience Surveys Reveal Gender-Biased Descriptions of Their Care Providers. *Journal of Medical Systems*, *45*(10), 90. https://doi.org/10.1007/s10916-021-01766-z

Hehman, E., Sutherland, C. A. M., Flake, J. K., & Slepian, M. L. (2017a). The unique contributions of perceiver and target characteristics in person perception. *Journal of Personality and Social Psychology*, *113*(4), 513–529. https://doi.org/10.1037/pspa0000090

Hehman, E., Xie, S. Y., Ofosu, E. K., & Nespoli, G. (2018). Assessing the point at which averages are stable: A tool illustrated in the context of person perception. OSF. https://doi.org/10.31234/osf.io/2n6jq

Hennein, R., Tineo, P., Bonumwezi, J., Gorman, H., Nguemeni Tiako, M. J., & Lowe, S. R. (2022). "They Wanted to Talk to a 'Real Doctor'": Predictors, Perpetrators, and Experiences of Racial and Ethnic Discrimination Among Healthcare Workers. *Journal of General Internal Medicine*, *37*(6), 1475–1483. https://doi.org/10.1007/s11606-021-07143-3

Herrmann, M. J., Würflein, H., Schreppel, T., Koehler, S., Mühlberger, A., Reif, A., Canli, T., Romanos, M., Jacob, C. P., Lesch, K.-P., & Fallgatter, A. J. (2009). Catechol-O-methyltransferase Val158Met genotype affects neural correlates of aversive stimuli processing. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(2), 168–172. https://doi.org/10.3758/CABN.9.2.168

Heybati, K., Chang, A., Mohamud, H., Satkunasivam, R., Coburn, N., Salles, A., Tsugawa, Y., Ikesu, R., Saka, N., Detsky, A. S., Ko, D. T., Ross, H., Mamas, M. A., Jerath, A., & Wallis, C. J. D. (2025). The association between physician sex and patient outcomes: A systematic review and meta-analysis. *BMC Health Services Research*, *25*(1), 93. https://doi.org/10.1186/s12913-025-12247-1

Hill, E. J. R., Bowman, K. A., Stalmeijer, R. E., Solomon, Y., & Dornan, T. (2014). Can I cut it? Medical students' perceptions of surgeons and surgical careers. *American Journal of Surgery*, *208*(5), 860–867. https://doi.org/10.1016/j.amjsurg.2014.04.016

Hillmer, K., Kappesser, J., & Hermann, C. (2021a). Pain modulation by your partner: An experimental investigation from a social-affective perspective. *PLOS ONE*, *16*(7), e0254069. https://doi.org/10.1371/journal.pone.0254069

Hillmer, K., Kappesser, J., & Hermann, C. (2024). Affective and social pain modulation in children—Experimental evidence using picture viewing. *PLOS ONE*, *19*(12), e0313636. https://doi.org/10.1371/journal.pone.0313636

Howe, L. C., Leibowitz, K. A., & Crum, A. J. (2019). When Your Doctor “Gets It” and “Gets You”: The Critical Role of Competence and Warmth in the Patient–Provider Interaction. *Frontiers in Psychiatry*, *10*. https://doi.org/10.3389/fpsyt.2019.00475

Inoue, H., Yamasue, H., Tochigi, M., Takei, K., Suga, M., Abe, O., Yamada, H., Rogers, M. A., Aoki, S., Sasaki, T., & Kasai, K. (2010). Effect of tryptophan hydroxylase-2 gene variants on amygdalar and hippocampal volumes. *Brain Research*, *1331*, 51–57. https://doi.org/10.1016/j.brainres.2010.03.057

Jaeger, B., Todorov, A. T., Evans, A. M., & van Beest, I. (2020). Can we reduce facial biases? Persistent effects of facial trustworthiness on sentencing decisions. *Journal of Experimental Social Psychology*, *90*, 104004. https://doi.org/10.1016/j.jesp.2020.104004

Jakubiak, B. K., & Feeney, B. C. (2016). Keep in touch: The effects of imagined touch support on stress and exploration. *Journal of Experimental Social Psychology*, *65*, 59–67. https://doi.org/10.1016/j.jesp.2016.04.001

Jaltare, K. P., Manresa, J. B., Niwa, S., & Torta, D. M. (2024). Verbal Support From a Stranger Reduces the Development of Mechanical Hypersensitivity: Behavioral and Neurophysiological Evidence. *The Journal of Pain*, *25*(10), 104599. https://doi.org/10.1016/j.jpain.2024.104599

Jaltare, K. P., Vanderijst, L., Karos, K., & Torta, D. M. (2023). The impact of the social context on the development of secondary hyperalgesia: An experimental study. *PAIN*, *164*(12), 2711. https://doi.org/10.1097/j.pain.0000000000002971

James, T. (2016). The Patient–Physician Clinical Encounter. In M. L. Martin, S. L. Heron, L. Moreno-Walton, & A. W. Jones (Eds.), *Diversity and Inclusion in Quality Patient Care* (pp. 69–81). Springer International Publishing. https://doi.org/10.1007/978-3-319-22840-2_7

Jenkins, R., White, D., Van Montfort, X., & Mike Burton, A. (2011). Variability in photos of the same face. *Cognition*, *121*(3), 313–323. https://doi.org/10.1016/j.cognition.2011.08.001

Jenkins, T. M. (2020). 1. Meet the Residents. In *Doctors’ Orders: The Making of Status Hierarchies in an Elite Profession* (pp. 33–72). Columbia University Press. https://www.degruyterbrill.com/document/doi/10.7312/jenk18934-005/html

Kanazawa, S. (2011). Intelligence and physical attractiveness. *Intelligence*, *39*(1), 7–14. https://doi.org/10.1016/j.intell.2010.11.003

Karos, K., Meulders, A., Goubert, L., & Vlaeyen, J. W. S. (2018). The Influence of Social Threat on Pain, Aggression, and Empathy in Women. *The Journal of Pain*, *19*(3), 291–300. https://doi.org/10.1016/j.jpain.2017.11.003

Karos, K., Meulders, A., Leyssen, T., & Vlaeyen, J. W. (2020). Freeze-like responses to pain in humans and its modulation by social context. *PeerJ*, *8*, e10094. https://doi.org/10.7717/peerj.10094

Kauff, M., Anslinger, J., Christ, O., Niemann, M., Geierhos, M., & Huster, L. (2022). Ethnic and gender-based prejudice towards medical doctors? The relationship between physicians' ethnicity, gender, and ratings on a physician rating website. *The Journal of Social Psychology*, *162*(5), 540–548. https://doi.org/10.1080/00224545.2021.1927944

Keefe, F. J., Lefebvre, J. C., Egert, J. R., Affleck, G., Sullivan, M. J., & Caldwell, D. S. (2000). The relationship of gender to pain, pain behavior, and disability in osteoarthritis patients: The role of catastrophizing. *Pain*, *87*(3), 325–334. https://doi.org/10.1016/S0304-3959(00)00296-7

Keller, L., & Reeve, H. K. (1998). Familiarity breeds cooperation. *Nature*, *394*(6689), 121–122. https://doi.org/10.1038/28031

Kessner, S., Sprenger, C., Wrobel, N., Wiech, K., & Bingel, U. (2013). Effect of oxytocin on placebo analgesia: A randomized study. *JAMA*, *310*(16), 1733–1735. https://doi.org/10.1001/jama.2013.277446

Kilford, E. J., Dumontheil, I., Wood, N. W., & Blakemore, S.-J. (2015). Influence of COMT genotype and affective distractors on the processing of self-generated thought. *Social Cognitive and Affective Neuroscience*, *10*(6), 777–782. https://doi.org/10.1093/scan/nsu118

Kleisner, K., Chvátalová, V., & Flegr, J. (2014). Perceived Intelligence Is Associated with Measured Intelligence in Men but Not Women. *PLOS ONE*, *9*(3), e81237. https://doi.org/10.1371/journal.pone.0081237

Koch, A. J., D'Mello, S. D., & Sackett, P. R. (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *Journal of Applied Psychology*, *100*(1), 128–161. https://doi.org/10.1037/a0036734

Krahé, C., Springer, A., Weinman, J. A., & Fotopoulou, A. (2013). The Social Modulation of Pain: Others as Predictive Signals of Salience – a Systematic Review. *Frontiers in Human Neuroscience*, *7*, 386. https://doi.org/10.3389/fnhum.2013.00386

Krupp, D. B., Debruine, L. M., & Barclay, P. (2008). A cue of kinship promotes cooperation for the public good. *Evolution and Human Behavior*, *29*(1), 49–55. https://doi.org/10.1016/j.evolhumbehav.2007.08.002

Ku, L., & Vichare, A. (2023). The Association of Racial and Ethnic Concordance in Primary Care with Patient Satisfaction and Experience of Care. *Journal of General Internal Medicine*, *38*(3), 727–732. https://doi.org/10.1007/s11606-022-07695-y

Laas, K., Kiive, E., Mäestu, J., Vaht, M., Veidebaum, T., & Harro, J. (2017). Nice guys: Homozygocity for the *TPH2* -703G/T (rs4570625) minor allele promotes low aggressiveness and low anxiety. *Journal of Affective Disorders*, *215*, 230–236. https://doi.org/10.1016/j.jad.2017.03.045

Laborde, S., Mosley, E., & Thayer, J. F. (2017). Heart Rate Variability and Cardiac Vagal Tone in Psychophysiological Research – Recommendations for Experiment Planning, Data Analysis, and Data Reporting. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00213

Landrigan, C. P., Rothschild, J. M., Cronin, J. W., Kaushal, R., Burdick, E., Katz, J. T., Lilly, C. M., Stone, P. H., Lockley, S. W., Bates, D. W., & Czeisler, C. A. (2004). Effect of reducing interns' work hours on serious medical errors in intensive care units. *The New England Journal of Medicine*, *351*(18), 1838–1848. https://doi.org/10.1056/NEJMoa041406

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry*, *44*(12), 1248–1263. https://doi.org/10.1016/s0006-3223(98)00275-3

Lavigne, G. J., Zucconi, M., Castronovo, V., Manzini, C., Veglia, F., Smirne, S., & Ferini-Strambi, L. (2001). Heart rate changes during sleep in response to experimental thermal (nociceptive) stimulations in healthy subjects. *Clinical Neurophysiology*, *112*(3), 532–535. https://doi.org/10.1016/S1388-2457(00)00558-7

Leknes, S., & Bastian, B. (2014). The Benefits of Pain. *Review of Philosophy and Psychology*, *5*(1), 57–70. https://doi.org/10.1007/s13164-014-0178-3

Leknes, S., Berna, C., Lee, M. C., Snyder, G. D., Biele, G., & Tracey, I. (2013). The importance of context: When relative relief renders pain pleasant. *PAIN*, *154*(3), 402. https://doi.org/10.1016/j.pain.2012.11.018

Leone, L., Perugini, M., Bagozzi, R. P., Pierro, A., & Mannetti, L. (2001). Construct validity and generalizability of the Carver–White behavioural inhibition system/behavioural activation system scales. *European Journal of Personality*, *15*(5), 373–390. https://doi.org/10.1002/per.415

Levy, A., Shalom, B. R., & Chalamish, M. (2025). A guide to similarity measures and their data science applications. *Journal of Big Data*, *12*(1), 188. https://doi.org/10.1186/s40537-025-01227-1

Li, J., Zhao, Y., Li, R., Broster, L. S., Zhou, C., & Yang, S. (2015). Association of Oxytocin Receptor Gene (OXTR) rs53576 Polymorphism with Sociality: A Meta-Analysis. *PLOS ONE*, *10*(6), e0131820. https://doi.org/10.1371/journal.pone.0131820

Ling, L., Luo, D., & SHE, G. (2019). Judging a book by its cover: The influence of physical attractiveness on the promotion of regional leaders. *Journal of Economic Behavior & Organization*, *158*, 1–14. https://doi.org/10.1016/j.jebo.2019.01.005

Livingston, R. W., & Pearce, N. A. (2009). The teddy-bear effect: Does having a baby face benefit black chief executive officers? *Psychological Science*, *20*(10), 1229–1236. https://doi.org/10.1111/j.1467-9280.2009.02431.x

LoParo, D., & Waldman, I. D. (2015). The oxytocin receptor gene (OXTR) is associated with autism spectrum disorder: A meta-analysis. *Molecular Psychiatry*, *20*(5), 640–646. https://doi.org/10.1038/mp.2014.77

Lotta, T., Vidgren, J., Tilgmann, C., Ulmanen, I., Melén, K., Julkunen, I., & Taskinen, J. (1995). Kinetics of human soluble and membrane-bound catechol O-methyltransferase: A revised mechanism and description of the thermolabile variant of the enzyme. *Biochemistry*, *34*(13), 4202–4210. https://doi.org/10.1021/bi00013a008

Lucht, M. J., Barnow, S., Sonnenfeld, C., Ulrich, I., Grabe, H. J., Schroeder, W., Völzke, H., Freyberger, H. J., John, U., Herrmann, F. H., Kroemer, H., & Rosskopf, D. (2013). Associations between the oxytocin receptor gene (OXTR) and "mind-reading" in humans—An exploratory study. *Nordic Journal of Psychiatry*, *67*(1), 15–21. https://doi.org/10.3109/08039488.2012.700731

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago Face Database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135.

Malloy, C. S., Hughes, C., & Cassidy, B. S. (2023). Perceiver and target partisanship shift facial trustworthiness effects on likability. *Scientific Reports*, *13*(1), 6130. https://doi.org/10.1038/s41598-023-33307-8

Mana, S. C., & Sasipraba, T. (2021). Research on Cosine Similarity and Pearson Correlation Based Recommendation Models. *Journal of Physics: Conference Series*, *1770*(1), 012014. https://doi.org/10.1088/1742-6596/1770/1/012014

March, E., van Dick, R., & Hernandez Bark, A. (2016). Current prescriptions of men and women in differing occupational gender roles. *Journal of Gender Studies*, *25*(6), 681–692. https://doi.org/10.1080/09589236.2015.1090303

Martinez, J. E., Funk, F., & Todorov, A. (2020). Quantifying idiosyncratic and shared contributions to judgment. *Behavior Research Methods*, *52*(4), 1428–1444. https://doi.org/10.3758/s13428-019-01323-0

Mattarozzi, K., Caponera, E., Russo, P. M., Colonnello, V., Bassetti, M., Farolfi, E., & Todorov, A. (2021). Pain and satisfaction: Healthcare providers' facial appearance matters. *Psychological Research*, *85*(4), 1706–1712. https://doi.org/10.1007/s00426-020-01330-3

Mattarozzi, K., Colonnello, V., De Gioia, F., & Todorov, A. (2017). I care, even after the first impression: Facial appearance-based evaluations in healthcare context. *Social Science & Medicine*, *182*, 68–72. https://doi.org/10.1016/j.socscimed.2017.04.011

Mattarozzi, K., Todorov, A., Marzocchi, M., Vicari, A., & Russo, P. M. (2015). Effects of gender and personality on first impression. *PloS One*, *10*(9), e0135529.

Mayo, L. M., Asratian, A., Lindé, J., Holm, L., Nätt, D., Augier, G., Stensson, N., Vecchiarelli, H. A., Balsevich, G., Aukema, R. J., Ghafouri, B., Spagnolo, P. A., Lee, F. S., Hill, M. N., & Heilig, M. (2020). Protective effects of elevated anandamide on stress and fear-related behaviors: Translational evidence from humans and mice. *Molecular Psychiatry*, *25*(5), 993–1005. https://doi.org/10.1038/s41380-018-0215-1

Mayo, L. M., Asratian, A., Lindé, J., Morena, M., Haataja, R., Hammar, V., Augier, G., Hill, M. N., & Heilig, M. (2020). Elevated Anandamide, Enhanced Recall of Fear Extinction, and Attenuated Stress Responses Following Inhibition of Fatty Acid Amide Hydrolase: A Randomized, Controlled Experimental Medicine Trial. *Biological Psychiatry*, *87*(6), 538–547. https://doi.org/10.1016/j.biopsych.2019.07.034

McCormick, F., Kadzielski, J., Landrigan, C. P., Evans, B., Herndon, J. H., & Rubash, H. E. (2012). Surgeon Fatigue: A Prospective Analysis of the Incidence, Risk, and Intervals of Predicted Fatigue-Related Impairment in Residents. *Archives of Surgery*, *147*(5), 430–435. https://doi.org/10.1001/archsurg.2012.84

McNeil, D. W., & Rainwater, A. J. (1998). Development of the Fear of Pain Questionnaire-III. *Journal of Behavioral Medicine*, *21*(4), 389–410. https://doi.org/10.1023/A:1018782831217

Mcquillan, J., & Hernandez, N. (2021). Real-Life Conundrums in the Struggle for Institutional Transformation. *Gender & Society*, *35*(3), 300–329. https://doi.org/10.1177/08912432211013147

Melchers, M., Montag, C., Felten, A., & Reuter, M. (2015). The oxytocin receptor gene and social perception. *Social Neuroscience*, *10*(4), 345–353. https://doi.org/10.1080/17470919.2015.1008646

Melchers, M., Montag, C., Markett, S., Niazy, N., Groß-Bölting, J., Zimmermann, J., & Reuter, M. (2017). The OXTR gene, implicit learning and social processing: Does empathy evolve from perceptual skills for details? *Behavioural Brain Research*, *329*, 35–40. https://doi.org/10.1016/j.bbr.2017.04.036

Mikulincer, M., & Shaver, P. R. (2010). *Attachment in Adulthood, First Edition: Structure, Dynamics, and Change*. Guilford Publications.

Mikulincer, M., Shaver, P. R., Sapir-Lavid, Y., & Avihou-Kanza, N. (2009). What's inside the minds of securely and insecurely attached people? The secure-base script and its associations with attachment-style dimensions. *Journal of Personality and Social Psychology*, *97*(4), 615–633. https://doi.org/10.1037/a0015649

Minozzi, W., & Woon, J. (2023). Electability salience can bias voting decisions. *Research & Politics*, *10*(1), 20531680231159074. https://doi.org/10.1177/20531680231159074

Mogilski, J. K., & Welling, L. L. M. (2018). The Relative Contribution of Jawbone and Cheekbone Prominence, Eyebrow Thickness, Eye Size, and Face Length to Evaluations of Facial Masculinity and Attractiveness: A Conjoint Data-Driven Approach. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02428

Monticone, M., Baiardi, P., Ferrari, S., Foti, C., Mugnai, R., Pillastrini, P., Rocca, B., & Vanti, C. (2012). Development of the Italian version of the Pain Catastrophising Scale (PCS-I): Cross-cultural adaptation, factor analysis, reliability, validity and sensitivity to change. *Quality of Life Research*, *21*(6), 1045–1050. https://doi.org/10.1007/s11136-011-0007-4

Montoya, R. M., Horton, R. S., & Kirchner, J. (2008). Is actual similarity necessary for attraction? A meta-analysis of actual and perceived similarity. *Journal of Social and Personal Relationships*, *25*(6), 889–922. https://doi.org/10.1177/0265407508096700

Montoya, R. M., Horton, R. S., Vevea, J. L., Citkowicz, M., & Lauber, E. A. (2017). A re-examination of the mere exposure effect: The influence of repeated exposure on recognition, familiarity, and liking. *Psychological Bulletin*, *143*(5), 459–498. https://doi.org/10.1037/bul0000085

Mosier, A. E., & Pietri, E. S. (2021). The link between stereotypical leader traits and perceived electability: Do Black women politicians face double jeopardy? *Analyses of Social Issues and Public Policy*, *21*(1), 374–397. https://doi.org/10.1111/asap.12244

Murphy, E., & Oesch, D. (2016). The Feminization of Occupations and Change in Wages: A Panel Analysis of Britain, Germany, and Switzerland. *Social Forces*, *94*(3), 1221–1255. https://doi.org/10.1093/sf/sov099

Murray, B., & McCrone, S. (2015). An integrative review of promoting trust in the patient–primary care provider relationship. *Journal of Advanced Nursing*, *71*(1), 3–23. https://doi.org/10.1111/jan.12502

Necka, E. A., Amir, C., Dildine, T. C., & Atlas, L. Y. (2021). Expectations about pain and analgesic treatment are shaped by medical providers' facial appearances: Evidence from five online clinical simulation experiments. *Social Science & Medicine*, *281*, 114091. https://doi.org/10.1016/j.socscimed.2021.114091

O'Brien, M. B., Copus, A., Johnson, J., & Inglehart, M. R. (2019). Examining the Impact of Dental Hygienists' Professional Appearance: Patients' and dental student providers' perspectives. *American Dental Hygienists' Association*, *93*(4), 33–43.

Oldmeadow, J. A., Sutherland, C. A. M., & Young, A. W. (2013). Facial Stereotype Visualization Through Image Averaging. *Social Psychological and Personality Science*, *4*(5), 615–623. https://doi.org/10.1177/1948550612469820

Olivola, C. Y., Tingley, D., & Todorov, A. (2018). Republican Voters Prefer Candidates Who Have Conservative-Looking Faces: New Evidence From Exit Polls. *Political Psychology*, *39*(5), 1157–1171. https://doi.org/10.1111/pops.12489

Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-Based Trait Inferences and Voting. *Journal of Nonverbal Behavior*, *34*(2), 83–110. https://doi.org/10.1007/s10919-009-0082-1

Oosterhof, N. N., & Todorov, A. (2008, August 12). *The functional basis of face evaluation*. Proceedings of the National Academy of Sciences. https://doi.org/10.1073/pnas.0805664105

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, *9*(1), 128–133. https://doi.org/10.1037/a0014520

Parak, J., & Korhonen, I. (2013). *Accuracy of Firstbeat Bodyguard 2 beat-to-beat heart rate monitor. White paper by Firstbeat Technologies Ltd.*

Parker, K. J., Garner, J. P., Libove, R. A., Hyde, S. A., Hornbeak, K. B., Carson, D. S., Liao, C.-P., Phillips, J. M., Hallmayer, J. F., & Hardan, A. Y. (2014). Plasma oxytocin concentrations and *OXTR* polymorphisms predict social impairments in children with and without autism spectrum disorder. *Proceedings of the National Academy of Sciences*, *111*(33), 12258–12263. https://doi.org/10.1073/pnas.1402236111

Pastor, M. C., Bradley, M. M., Löw, A., Versace, F., Moltó, J., & Lang, P. J. (2008). Affective picture perception: Emotion, context, and the late positive potential. *Brain Research*, *1189*, 145–151. https://doi.org/10.1016/j.brainres.2007.10.072

Patel, R., & Wyse, E. (2025). ‘Bad manners’—Refocusing on ‘bedside manner’ as a core professional behaviour for doctors. *Postgraduate Medical Journal*, qgaf160. https://doi.org/10.1093/postmj/qgaf160

Peciña, M., Martínez-Jauand, M., Hodgkinson, C., Stohler, C. S., Goldman, D., & Zubieta, J. K. (2014). FAAH selectively influences placebo effects. *Molecular Psychiatry*, *19*(3), 385–391. https://doi.org/10.1038/mp.2013.124

Pedrabissi, L., & Santinello, M. (1989). *Manuale per la nuova versione italiana dello S.T.A.I. - forma Y*. https://www.research.unipd.it/handle/11577/2523749

Penttilä, J., Helminen, A., Jartti, T., Kuusela, T., Huikuri, H. V., Tulppo, M. P., Coffeng, R., & Scheinin, H. (2001). Time domain, geometrical and frequency domain analysis of cardiac vagal outflow: Effects of various respiratory patterns. *Clinical Physiology (Oxford, England)*, *21*(3), 365–376. https://doi.org/10.1046/j.1365-2281.2001.00337.x

Peskin, M., & Newell, F. N. (2004). Familiarity breeds attraction: Effects of exposure on the attractiveness of typical and distinctive faces. *Perception*, *33*(2), 147–157. https://doi.org/10.1068/p5028

Peterson, J., Griffiths, T., Uddenberg, S., Todorov, A., & Suchow, J. W. (2022). Deep models of superficial face judgments. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4041458

Petrilli, C. M., Saint, S., Jennings, J. J., Caruso, A., Kuhn, L., Snyder, A., & Chopra, V. (2018). *Understanding patient preference for physician attire: A cross-sectional observational study of 10 academic medical centres in the USA*. https://doi.org/10.1136/bmjopen-2017-021239

Porter, S., ten Brinke, L., & Gustaw, C. (2010). Dangerous decisions: The impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime & Law*, *16*(6), 477–491. https://doi.org/10.1080/10683160902926141

Price, D. D. (2000). Psychological and neural mechanisms of the affective dimension of pain. *Science (New York, N.Y.)*, *288*(5472), 1769–1772. https://doi.org/10.1126/science.288.5472.1769

Raja, S. N., Carr, D. B., Cohen, M., Finnerup, N. B., Flor, H., Gibson, S., Keefe, F. J., Mogil, J. S., Ringkamp, M., Sluka, K. A., Song, X.-J., Stevens, B., Sullivan, M. D., Tutelman, P. R., Ushida, T., & Vader, K. (2020). The revised International Association for the Study of Pain definition of pain:

Concepts, challenges, and compromises. *PAIN*, *161*(9), 1976. https://doi.org/10.1097/j.pain.0000000000001939

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Re, D. E., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2013). Facial Cues to Perceived Height Influence Leadership Choices in Simulated War and Peace Contexts. *Evolutionary Psychology*, *11*(1), 89–103. https://doi.org/10.1177/147470491301100109

Reddan, M. C., Young, H., Falkner, J., López-Solà, M., & Wager, T. D. (2020). Touch and social support influence interpersonal synchrony and pain. *Social Cognitive and Affective Neuroscience*, *15*(10), 1064–1075. https://doi.org/10.1093/scan/nsaa048

Redman, T., & Snape, E. (2002). Ageism in Teaching: Stereotypical Beliefs and Discriminatory Attitudes towards the Over-50s. *Work, Employment and Society*, *16*(2), 355–371. https://doi.org/10.1177/095001702400426884

Rehman, S. U., Nietert, P. J., Cope, D. W., & Kilpatrick, A. O. (2005a). What to wear today? Effect of doctor's attire on the trust and confidence of patients. *The American Journal of Medicine*, *118*(11), 1279–1286. https://doi.org/10.1016/j.amjmed.2005.04.026

Reuter, M., Frenzel, C., Walter, N. T., Markett, S., & Montag, C. (2011). Investigating the genetic basis of altruism: The role of the COMT Val158Met polymorphism. *Social Cognitive and Affective Neuroscience*, *6*(5), 662–668. https://doi.org/10.1093/scan/nsq083

Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable Facial Configurations Affect Strategic Choices in Trust Games with or without Information about Past Behavior. *PLOS ONE*, *7*(3), e34293. https://doi.org/10.1371/journal.pone.0034293

Riello, M., Cecchini, M. P., Zanini, A., Di Chiappari, M., Tinazzi, M., & Fiorio, M. (2019). Perception of phasic pain is modulated by smell and taste. *European Journal of Pain*, *23*(10), 1790–1800. https://doi.org/10.1002/ejp.1453

Rizvi, M. A. K., & Hossain, M. Z. (2017). Relationship Between Religious Belief and Happiness: A Systematic Literature Review. *Journal of Religion and Health*, *56*(5), 1561–1582. https://doi.org/10.1007/s10943-016-0332-6

Roberts, M. H., Klatzkin, R. R., & Mechlin, B. (2015). Social Support Attenuates Physiological Stress Responses and Experimental Pain Sensitivity to Cold Pressor Pain. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, *49*(4), 557–569. https://doi.org/10.1007/s12160-015-9686-3

Roter, D. L., & Hall, J. A. (2004). Physician Gender and Patient-Centered Communication: A Critical Review of Empirical Research. *Annual Review of Public Health*, *25*(Volume 25, 2004), 497–519. https://doi.org/10.1146/annurev.publhealth.25.101802.123134

Rougier, M., De Houwer, J., Richetin, J., Hughes, S., & Perugini, M. (2023). From Halo to Conditioning and Back Again: Exploring the Links Between Impression Formation and Learning. *Collabra: Psychology*, *9*(1), 84560. https://doi.org/10.1525/collabra.84560

Rowland, P. A., Coe, N. P. W., Burchard, K. W., & Pricolo, V. E. (2005). Factors affecting the professional image of physicians. *Current Surgery*, *62*(2), 214–219. https://doi.org/10.1016/j.cursur.2004.08.008

Saha, S., & Beach, M. C. (2020). Impact of Physician Race on Patient Decision-Making and Ratings of Physicians: A Randomized Experiment Using Video Vignettes. *Journal of General Internal Medicine*, *35*(4), 1084–1091. https://doi.org/10.1007/s11606-020-05646-z

Sambo, C. F., Howard, M., Kopelman, M., Williams, S., & Fotopoulou, A. (2010). Knowing you care: Effects of perceived empathy and attachment style on pain perception. *Pain*, *151*(3), 687–693. https://doi.org/10.1016/j.pain.2010.08.035

Schäfer, G., Prkachin, K. M., Kaseweter, K. A., & Williams, A. C. de C. (2016). Health care providers' judgments in chronic pain: The influence of gender and trustworthiness. *PAIN*, *157*(8), 1618. https://doi.org/10.1097/j.pain.0000000000000536

Schattner, A., Rudin, D., & Jellin, N. (2004). Good physicians from the perspective of their patients. *BMC Health Services Research*, *4*(1), 26. https://doi.org/10.1186/1472-6963-4-26

Scheuch, K., Lautenschlager, M., Grohmann, M., Stahlberg, S., Kirchheiner, J., Zill, P., Heinz, A., Walther, D. J., & Priller, J. (2007). Characterization of a Functional Promoter Polymorphism of the Human Tryptophan Hydroxylase 2 Gene in Serotonergic Raphe Neurons. *Biological Psychiatry*, *62*(11), 1288–1294. https://doi.org/10.1016/j.biopsych.2007.01.015

Schiller, D., & Delgado, M. R. (2010). Overlapping neural systems mediating extinction, reversal and regulation of fear. *Trends in Cognitive Sciences*, *14*(6), 268–276. https://doi.org/10.1016/j.tics.2010.04.002

Schmahl, C., Ludäscher, P., Greffrath, W., Kraus, A., Valerius, G., Schulze, T. G., Treutlein, J., Rietschel, M., Smolka, M. N., & Bohus, M. (2012). *PLOS ONE*, *7*(1), e23658. https://doi.org/10.1371/journal.pone.0023658

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime (version 2.0). Pittsburgh, PA: Psychology Software Tools Inc*. Retrieved 18/04/2019 from https://pstnet. com/products/e-prime.

Serrano, J. M., Banks, J. B., Fagan, T. J., & Tartar, J. L. (2019). The influence of Val158Met COMT on physiological stress responsivity. *Stress*, *22*(2), 276–279. https://doi.org/10.1080/10253890.2018.1553949

Shapiro, C. M., Auch, C., Reimer, M., Kayumov, L., Heslegrave, R., Huterer, N., Driver, H., & Devins, G. M. (2006). A new approach to the construct of alertness. *Journal of Psychosomatic Research*, *60*(6), 595–603. https://doi.org/10.1016/j.jpsychores.2006.04.012

Shen, M. J., Peterson, E. B., Costas-Muñiz, R., Hernandez, M. H., Jewell, S. T., Matsoukas, K., & Bylund, C. L. (2018). The Effects of Race and Racial Concordance on Patient-Physician Communication: A Systematic Review of the Literature. *Journal of Racial and Ethnic Health Disparities*, *5*(1), 117–140. https://doi.org/10.1007/s40615-017-0350-4

Siedlecka, M., Spychała, N., Łukowska, M., Wiercioch, K., & Wierzchoń, M. (2018). Rubber Hand Illusion Increases Pain Caused by Electric Stimuli. *The Journal of Pain*, *19*(1), 35–45. https://doi.org/10.1016/j.jpain.2017.08.005

Simon, D. H., & Lewis-Liu, T. (2024). *Assessing Patients' Preference for Physician-Patient Race Concordance: Evidence from Two Experiments* (SSRN Scholarly Paper No. 5077253). Social Science Research Network. https://doi.org/10.2139/ssrn.5077253

Sipe, J. C., Chiang, K., Gerber, A. L., Beutler, E., & Cravatt, B. F. (2002). A missense mutation in human fatty acid amide hydrolase associated with problem drug use. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(12), 8394–8399. https://doi.org/10.1073/pnas.082235799

Skuse, D. H., & Gallagher, L. (2011). Genetic Influences on Social Cognition. *Pediatric Research*, *69*(8), 85–91. https://doi.org/10.1203/PDR.0b013e318212f562

Smolka, M. N., Schumann, G., Wrase, J., Grüsser, S. M., Flor, H., Mann, K., Braus, D. F., Goldman, D., Büchel, C., & Heinz, A. (2005). Catechol-O-methyltransferase val158met genotype affects processing of emotional stimuli in the amygdala and prefrontal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(4), 836–842. https://doi.org/10.1523/JNEUROSCI.1792-04.2005

Soeiro-de-Souza, M. G., Bio, D. S., David, D. P., Rodrigues dos Santos, D., Kerr, D. S., Gattaz, W. F., Machado-Vieira, R., & Moreno, R. A. (2012). COMT Met (158) modulates facial emotion recognition in bipolar I disorder mood episodes. *Journal of Affective Disorders*, *136*(3), 370–376. https://doi.org/10.1016/j.jad.2011.11.021

Sofer, C., Dotsch, R., Wigboldus, D. H. J., & Todorov, A. (2015). What is typical is good: The influence of face typicality on perceived trustworthiness. *Psychological Science*, *26*(1), 39–47. https://doi.org/10.1177/0956797614554955

Somnath, S., & Beach, M. C. (2020). Impact of Physician Race on Patient Decision-Making and Ratings of Physicians: A Randomized Experiment Using Video Vignettes. *Journal of General Internal Medicine*, *35*(4), 1084–1091. https://doi.org/10.1007/s11606-020-05646-z

Spagnolo, P. A., Ramchandani, V. A., Schwandt, M. L., Kwako, L. E., George, D. T., Mayo, L. M., Hillard, C. J., & Heilig, M. (2016). FAAH Gene Variation Moderates Stress Response and Symptom Severity in Patients with Posttraumatic Stress Disorder and Comorbid Alcohol Dependence. *Alcoholism: Clinical and Experimental Research*, *40*(11), 2426–2434. https://doi.org/10.1111/acer.13210

Spielberger, C. D. (1989a). *State-trait anxiety inventory: Bibliography, 2nd edn. Palo Alto*. CA: Consulting Psychologists Press.

Stavely, T., Salhi, B. A., Lall, M. D., & Zeidan, A. (2022). “I just assume they don’t know that I’m the doctor”: Gender bias and professional identity development of women residents. *AEM Education and Training*, *6*(2), e10735. https://doi.org/10.1002/aet2.10735

Stoker, J. I., Garretsen, H., & Spreeuwers, L. J. (2016). The Facial Appearance of CEOs: Faces Signal Selection but Not Performance. *PLOS ONE*, *11*(7), e0159950. https://doi.org/10.1371/journal.pone.0159950

Stolier, R. M., Hehman, E., & Freeman, J. B. (2020). Trait knowledge forms a common structure across social cognition. *Nature Human Behaviour*, *4*(4), 361–371. https://doi.org/10.1038/s41562-019-0800-6

Stone, W. J., & Rapoport, R. B. (1994). Candidate Perception among Nomination Activists: A New Look at the Moderation Hypothesis. *The Journal of Politics*, *56*(4), 1034–1052. https://doi.org/10.2307/2132072

Street, R. L., O’Malley, K. J., Cooper, L. A., & Haidet, P. (2008a). Understanding Concordance in Patient-Physician Relationships: Personal and Ethnic Dimensions of Shared Identity. *The Annals of Family Medicine*, *6*(3), 198–205. https://doi.org/10.1370/afm.821

Sullivan, M. D., & de C Williams, A. C. (2025). The social nature of human pain. *PAIN*, *166*(1), 20. https://doi.org/10.1097/j.pain.0000000000003250

Sullivan, M. J. L. (2012). The communal coping model of pain catastrophising: Clinical and research implications. *Canadian Psychology / Psychologie Canadienne*, *53*(1), 32–41. https://doi.org/10.1037/a0026726

Sullivan, M. J. L., Bishop, S. R., & Pivik, J. (1995). The Pain Catastrophizing Scale: Development and validation. *Psychological Assessment*, *7*(4), 524–532. https://doi.org/10.1037/1040-3590.7.4.524

Takeshita, J., Wang, S., Loren, A. W., Mitra, N., Shults, J., Shin, D. B., & Sawinski, D. L. (2020). Association of Racial/Ethnic and Gender Concordance Between Patients and Physicians With Patient Experience Ratings. *JAMA Network Open*, *3*(11), e2024583. https://doi.org/10.1001/jamanetworkopen.2020.24583

Tang, J., & Gibson, S. J. (2005). A psychophysical evaluation of the relationship between trait anxiety, pain perception, and induced state anxiety. *The Journal of Pain*, *6*(9), 612–619. https://doi.org/10.1016/j.jpain.2005.03.009

Tarvainen, M. P., Niskanen, J.-P., Lipponen, J. A., Ranta-Aho, P. O., & Karjalainen, P. A. (2014). Kubios HRV--heart rate variability analysis software. *Computer Methods and Programs in Biomedicine*, *113*(1), 210–220. https://doi.org/10.1016/j.cmpb.2013.07.024

Tarvainen, M. P., Ranta-Aho, P. O., & Karjalainen, P. A. (2002). An advanced detrending method with application to HRV analysis. *IEEE Transactions on Bio-Medical Engineering*, *49*(2), 172–175. https://doi.org/10.1109/10.979357

Tedeschi, E., Weber, J., Prévost, C., Mischel, W., & Mobbs, D. (2015). Inferences of Others' Competence Reduces Anticipation of Pain When under Threat. *Journal of Cognitive Neuroscience*, *27*(10), 2071–2078. https://doi.org/10.1162/jocn_a_00843

Tella, M. D., Ghiggia, A., Testa, S., Castelli, L., & Adenzato, M. (2019). The Fear of Pain Questionnaire: Factor structure, validity and reliability of the Italian translation. *PLOS ONE*, *14*(1), e0210757. https://doi.org/10.1371/journal.pone.0210757

Thompson, R. J., Parker, K. J., Hallmayer, J. F., Waugh, C. E., & Gotlib, I. H. (2011a). Oxytocin receptor gene polymorphism (rs2254298) interacts with familial risk for psychopathology to predict symptoms of depression and anxiety in adolescent girls. *Psychoneuroendocrinology*, *36*(1), 144–147. https://doi.org/10.1016/j.psyneuen.2010.07.003

Tingley, D. (2014). Face-off: Facial features and strategic choice. *Political Psychology*, *35*(1), 35–55. https://doi.org/10.1111/pops.12041

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science (New York, N.Y.)*, *308*(5728), 1623–1626. https://doi.org/10.1126/science.1110589

Todorov, A., & Oh, D. (2021). The structure and perceptual basis of social judgments from faces. In *Advances in Experimental Social Psychology* (Vol. 63, pp. 189–245). Elsevier. https://doi.org/10.1016/bs.aesp.2020.11.004

Todorov, A., Oh, D., Uddenberg, S., & Albohn, D. N. (2025). Face evaluation: Findings, methods, and challenges. *Annals of the New York Academy of Sciences*, *1545*(1), 28–37. https://doi.org/10.1111/nyas.15293

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annual Review of Psychology*, *66*(Volume 66, 2015), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Tousignant-Laflamme, Y., Rainville, P., & Marchand, S. (2005). Establishing a Link Between Heart Rate and Pain in Healthy Subjects: A Gender Effect. *The Journal of Pain*, *6*(6), 341–347. https://doi.org/10.1016/j.jpain.2005.01.351

Tran, T. B., Raoof, M., Melstrom, L., Kyulo, N., Shaikh, Z., Jones, V. C., Erhunmwunsee, L., Fong, Y., & Warner, S. G. (2021). Racial and Ethnic Bias Impact Perceptions of Surgeon Communication. *Annals of Surgery*, *274*(4), 597. https://doi.org/10.1097/SLA.0000000000005060

Treister, R., Kliger, M., Zuckerman, G., Aryeh, I. G., & Eisenberg, E. (2012a). Differentiating between heat pain intensities: The combined effect of multiple autonomic parameters. *PAIN®*, *153*(9), 1807–1814. https://doi.org/10.1016/j.pain.2012.04.008

Treister, R., Kliger, M., Zuckerman, G., Aryeh, I. G., & Eisenberg, E. (2012b). Differentiating between heat pain intensities: The combined effect of multiple autonomic parameters. *PAIN®*, *153*(9), 1807–1814. https://doi.org/10.1016/j.pain.2012.04.008

Trost, Z., Scott, W., Lange, J. m., Manganelli, L., Bernier, E., & Sullivan, M. j. (2014). An experimental investigation of the effect of a justice violation on pain experience and expression among individuals with high and low just world beliefs. *European Journal of Pain*, *18*(3), 415–423. https://doi.org/10.1002/j.1532-2149.2013.00375.x

Tsugawa, Y., Jena, A. B., Figueroa, J. F., Orav, E. J., Blumenthal, D. M., & Jha, A. K. (2017). Comparison of Hospital Mortality and Readmission Rates for Medicare Patients Treated by Male vs Female Physicians. *JAMA Internal Medicine*, *177*(2), 206–213. https://doi.org/10.1001/jamainternmed.2016.7875

van Dulmen, A. M., Verhaak, P. F., & Bilo, H. J. (1997). Shifts in doctor-patient communication during a series of outpatient consultations in non-insulin-dependent diabetes mellitus. *Patient Education and Counseling*, *30*(3), 227–237. https://doi.org/10.1016/s0738-3991(96)00965-2

van 't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition*, *108*(3), 796–803. https://doi.org/10.1016/j.cognition.2008.07.002

Vargas-Alarcón, G., Fragoso, J.-M., Cruz-Robles, D., Vargas, A., Vargas, A., Lao-Villadóniga, J.-I., García-Fructuoso, F., Ramos-Kuri, M., Hernández, F., Springall, R., Bojalil, R., Vallejo, M., & Martínez-Lavín, M. (2007). Catechol-O-methyltransferase gene haplotypes in Mexican and Spanish patients with fibromyalgia. *Arthritis Research & Therapy*, *9*(5), R110. https://doi.org/10.1186/ar2316

Vella, K., Koren, C. J., & Johnson, D. (2017). The Impact of Agency and Familiarity in Cooperative Multiplayer Games. *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 423–434. https://doi.org/10.1145/3116595.3116622

Vernon, R. J. W., Sutherland, C. A. M., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences*, *111*(32), E3353–E3361. https://doi.org/10.1073/pnas.1409860111

Verosky, S. C., & Todorov, A. (2010). Generalization of Affective Learning About Faces to Perceptually Similar Faces. *Psychological Science*, *21*(6), 779–785. https://doi.org/10.1177/0956797610371965

Villemure, C., & Bushnell, M. C. (2002). Cognitive modulation of pain: How do attention and emotion influence pain processing? *Pain*, *95*(3), 195–199. https://doi.org/10.1016/S0304-3959(02)00007-6

Villemure, C., & Bushnell, M. C. (2009). Mood influences supraspinal pain processing separately from attention. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(3), 705–715. https://doi.org/10.1523/JNEUROSCI.3822-08.2009

Vlaeyen, J. W. S., Hanssen, M., Goubert, L., Vervoort, T., Peters, M., van Breukelen, G., Sullivan, M. J. L., & Morley, S. (2009). Threat of pain influences social context effects on verbal pain report

and facial expression. *Behaviour Research and Therapy*, *47*(9), 774–782. https://doi.org/10.1016/j.brat.2009.05.008

Waider, J., Araragi, N., Gutknecht, L., & Lesch, K.-P. (2011). Tryptophan hydroxylase-2 (TPH2) in disorders of cognitive control and emotion regulation: A perspective. *Psychoneuroendocrinology*, *36*(3), 393–405. https://doi.org/10.1016/j.psyneuen.2010.12.012

Wallis, C. J., Ravi, B., Coburn, N., Nam, R. K., Detsky, A. S., & Satkunasivam, R. (2017). Comparison of postoperative outcomes among patients treated by male and female surgeons: A population based matched cohort study. *BMJ*, *359*, j4366. https://doi.org/10.1136/bmj.j4366

Watson, D., Clark, L. A., & Tellegen, A. (1988a). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Wei, X., Yu, S., & Li, C. (Victor). (2024). Influence of Physical Attractiveness and Gender on Patient Preferences in Digital Doctor Consultations: Experimental Study. *Journal of Medical Internet Research*, *26*(1), e46551. https://doi.org/10.2196/46551

Winkielman, P., Schwarz, N., Fazendeiro, T. A., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189–217). Lawrence Erlbaum Associates Publishers.

Wu, D. C., Kim, H. S., & Collins, N. L. (2021). Perceived responsiveness across cultures: The role of cultural fit in social support use. *Social and Personality Psychology Compass*, *15*(9), e12634. https://doi.org/10.1111/spc3.12634

Yao, S., Lix, L., Teare, G., Evans, C., & Blackburn, D. (2022). The Impact of Age and Sex Concordance Between Patients and Physicians on Medication Adherence: A Population-Based Study. *Patient Preference and Adherence*, *16*, 169–178. https://doi.org/10.2147/PPA.S340573

Zebrowitz, L. A., Bronstad, P. M., & Lee, H. K. (2007). The Contribution of Face Familiarity to Ingroup Favoritism and Stereotyping. *Social Cognition*, *25*(2), 306–338. https://doi.org/10.1521/soco.2007.25.2.306

Zebrowitz, L. A., Fellous, J.-M., Mignault, A., & Andreoletti, C. (2003). Trait Impressions as Overgeneralized Responses to Adaptively Significant Facial Qualities: Evidence from Connectionist

Modeling. *Personality and Social Psychology Review*, *7*(3), 194–215. https://doi.org/10.1207/S15327957PSPR0703_01

Zebrowitz, L. A., & Montepare, J. (2006). The Ecological Approach to Person Perception: Evolutionary Roots and Contemporary Offshoots. In *Evolution and Social Psychology*. Psychology Press.

Zebrowitz, L. A., & Montepare, J. M. (1992). Impressions of babyfaced individuals across the life span. *Developmental Psychology*, *28*(6), 1143–1152. https://doi.org/10.1037/0012-1649.28.6.1143

Zebrowitz, L. A., & Montepare, J. M. (2008). Social Psychological Face Perception: Why Appearance Matters. *Social and Personality Psychology Compass*, *2*(3), 1497–1517. https://doi.org/10.1111/j.1751-9004.2008.00109.x

Zebrowitz, L. A., Wang, R., Bronstad, P. M., Eisenberg, D., Undurraga, E., Reyes-García, V., & Godoy, R. (2012). First Impressions From Faces Among U.S. and Culturally Isolated Tsimane' People in the Bolivian Rainforest. *Journal of Cross-Cultural Psychology*, *43*(1), 119–134. https://doi.org/10.1177/0022022111411386

Zebrowitz, L. A., White, B., & Wieneke, K. (2008). Mere Exposure and Racial Prejudice: Exposure to Other-Race Faces Increases Liking for Strangers of that Race. *Social Cognition*, *26*(3), 259–275. https://doi.org/10.1521/soco.2008.26.3.259

Zhang, X., Wang, Y., Xiao, Q., & Wang, J. (2024). The impact of doctors' facial attractiveness on users' choices in online health communities: A stereotype content and social role perspective. *Decision Support Systems*, *182*, 114246. https://doi.org/10.1016/j.dss.2024.114246

8 Supplementary materials

8.1 Chapter 4: additional results

8.1.1 Exploratory analyses: the role of sex differences

Analyses including sex as a control factor evidenced that the observed effect of Condition was not driven by sex differences neither for pain intensity ($F(1,138)=1.80$, $p=.18$, $R^2_c=.75$, $R^2_m=.01$) nor unpleasantness ($F(1,138)=.02$, $p=.87$, $R^2_c=.72$, $R^2_m=.002$).

The interaction between Sex and Condition was significant for both variables considered (pain intensity: $F(2,3216)=4.14$, $p<.05$, $R^2_c=.75$, $R^2_m=.01$; unpleasantness: $F(2,3216)=13.52$, $p<.001$, $R^2_c=.72$, $R^2_m=.004$).

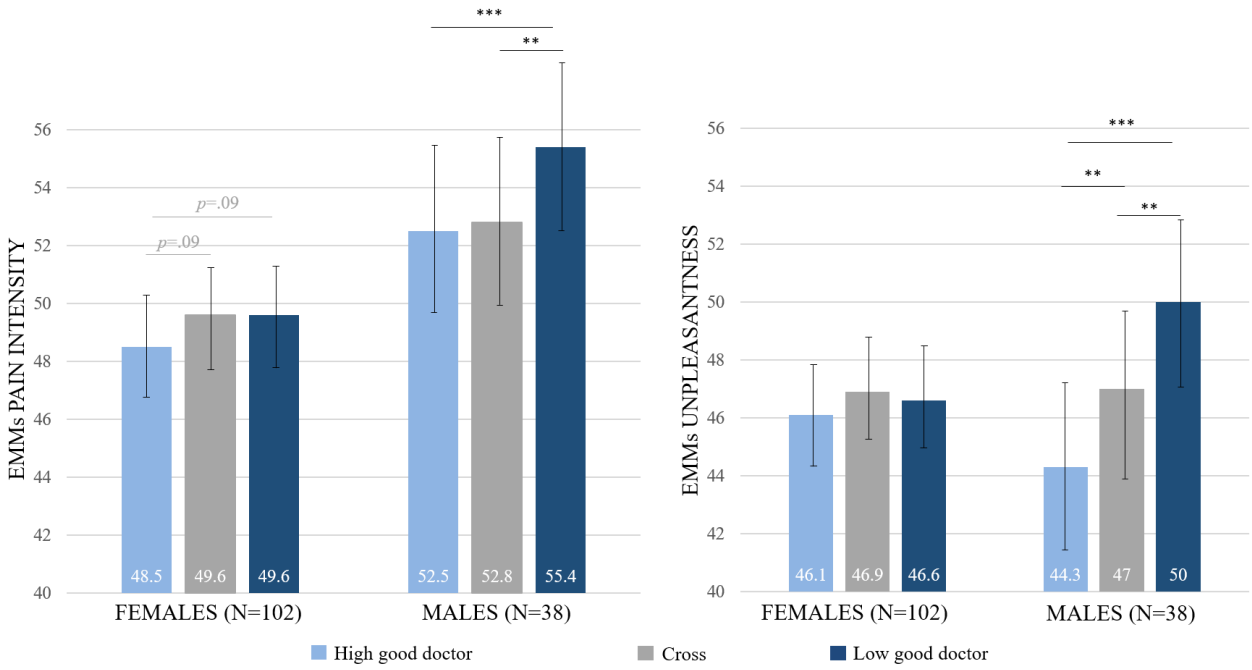
Specifically, we found that male participants reported significantly lower pain intensity in the "high good doctor" (*EMM*=52.50, 95% *CI*[47.00,58.00], $p<.001$) and fixation cross (*EMM*= 52.80, 95% *CI*[47.30, 58.30],$p<.01$) conditions compared to the "low good doctor" condition (*EMM*=55.40, 95% *CI*[49.90, 60.90]). In contrast, although female participants reported lower pain intensity in the "high good doctor" (*EMM*=48.50, 95% *CI*[45.20,51.90]) condition compared to both the "low good doctor" (*EMM*=49.60, *CI*[46.30,52.90], $p=.09$) and fixation cross (*EMM*=49.60, *CI*[46.30,52.90], $p=.09$) conditions, these differences did not reach significance.

Results on unpleasantness mirrored these findings. Analyses revealed that male participants reported significant lower unpleasantness in the "high good doctor" (*EMM*= 44.30, 95% *CI*[38.6, 49.9]) compared to both the "low good doctor" (*EMM*=50.0, 95% *CI*[44.30, 55.70],$p<.001$) and fixation cross (*EMM*=47.0, 95% *CI*[41.3, 52.6],$p<.01$) conditions. Moreover, a significant difference also emerged between the "low good doctor" and the fixation cross conditions ($p<.01$), with the former being associated with the highest unpleasantness ratings. No significant differences emerged in female participants across conditions.

Results are displayed in Figure S8.1.

Figure S8.1

Estimated marginal means (EMMs) of pain intensity and unpleasantness across the "high good doctor", "low good doctor" and fixation cross conditions in female and male participants.



Note. Error bars represent the standard error of the EMMs**.** **p*<.05, ** *p*<.01, *** *p*<.001

The complete output of the linear mixed effect models and the corresponding Bonferroni- adjusted pairwise comparisons are illustrated in Table S8.1 and S8.2.

Table S8.1

Linear- mixed effects regression table testing the Condition x Participant Sex interaction on pain intensity and unpleasantness.

	Pain Intenstiy				
Fixed effects	**Estimate (β)**	**SE**	**df**	***t* value**	***p***
Intercept (Fixation Cross x Females)	49.59	1.70	145.90	29.10	<.001
High vs Fixation Cross	-1.05	0.49	3216.00	-2.16	<.05
Low vs Fixation Cross	0.006	0.49	3216.00	0.01	.99
Males	3.19	3.27	145.90	0.98	.33

High x Males vs Fixation Cross x Females	0.78	0.94	3216.00	0.84	.40
Low x Males vs Fixation Cross x Females	2.62	0.94	3216.00	2.80	<.01

Random effects	**Variance**	**SD**
ID (Intercept)	284.11	16.86
Residual	97.13	9.86

	Unpleasantness				
Fixed effects	**Estimate (β)**	**SE**	**df**	***t* value**	***p***
Intercept (Fixation Cross x Females)	46.93	1.76	146.91	26.60	<.001
High vs Fixation Cross	-0.80	0.54	3216.00	-1.50	.13
Low vs Fixation Cross	0.34	0.54	3216.00	-0.63	.53
Males	0.02	3.39	146.91	0.01	.99
High x Males vs Fixation Cross x Females	-1.89	1.03	3216.00	-1.83	.07
Low x Males vs Fixation Cross x Females	3.40	1.03	3216.00	3.30	<.001

Random effects	**Variance**	**SD**
ID (Intercept)	302.80	17.40
Residual	117.50	10.84

Note. Baseline=Fixation cross in females. Negative coefficients indicate lower pain intensity and unpleasantness ratings compared to the baseline condition. The random intercept captures differences in pain intensity and unpleasantness ratings in the fixation cross condition for female participants.

Table S8.2

Bonferroni adjusted pairwise comparisons (EMMs) for Condition x Participant Sex interaction on pain intensity and unpleasantness.

	Pain Intenstiy				
Contrasts (Males)	**Estimate**	**SE**	**df**	**z ratio**	***p* (adjusted)**
Cross - High	0.27	0.80	3216	0.34	1.0
Cross - Low	-2.63	0.80	3216	-3.29	<.01
High - Low	-2.9	0.80	3216	-3.63	.09
Contrasts (Females)					
Cross - High	1.05	0.49	3216	2.16	.09
Cross - Low	-0.01	0.49	3216	-0.01	1.0
High - Low	-1.06	0.49	3216	-2.17	.08

	Unpleasantness				
Contrasts (Males)	**Estimate**	**SE**	**df**	**z ratio**	***p* (adjusted)**
Cross - High	2.69	0.88	3216	3.06	<.01
Cross - Low	-3.06	0.88	3216	-3.48	<.01
High - Low	-5.75	0.88	3216	-6.54	<.001
Contrasts (Females)					
Cross - High	0.81	0.54	3216	1.50	.40
Cross - Low	0.34	0.54	3216	0.63	1.00
High - Low	-0.47	0.54	3216	-0.87	1.00

Finally, the three-way interaction between Condition, participant's sex and the sex of facial stimuli did not lead to significant results for either pain intensity ($F(2094,1)=2.25$, $p=.13$, $R^2_c=.78$,$R^2_m=.02$) nor unpleasantness ($F(2094,1)=0.55$, $p=.46$, $R^2_c=.75$, $R^2_m=.01$).

The complete output of the linear mixed-effects model for the Condition x Participant Sex x Stimulus Sex interaction is displayed in Table S8.3

Table S8.3

Linear- mixed effects regression table testing the Condition x Participant Sex interaction on pain intensity and unpleasantness.

	Pain Intenstiy				
Fixed effects	**Estimate (β)**	**SE**	**df**	***t* value**	***p***
Intercept (High x Females x Female Face)	48.14	1.77	152.64	27.18	<.001
Low vs High	1.70	0.64	2094.00	2.64	<.01
Males vs Females	5.02	3.40	152.63	1.48	.14
Male Faces vs Female Faces	0.80	0.64	2094.00	1.24	.21
Low x Males vs High x Females	0.53	1.23	2094.00	0.43	.66
Low Male Faces vs High Female Faces	-1.27	0.91	2094.00	-1.40	.16
Males x Male Faces vs Females x Female Faces	-2.10	1.23	2094.00	-1.70	.09
Low x Males x Male Faces vs High x Females x Female Faces	2.61	1.74	2094.00	1.50	.13

Random effects	**Variance**	**SD**
ID (Intercept)	299.05	17.29
Residual	84.05	9.17

	Unpleasantness				
Fixed effects	**Estimate (β)**	**SE**	**df**	***t* value**	***p***
Intercept (High x Females x Female Face)	45.63	1.83	155.25	24.89	<.001
Low vs High	1.09	0.72	2094.00	1.51	.13
Males vs Females	-1.30	3.52	155.35	-0.37	.71
Male Faces vs Female Faces	1.00	0.72	2094.00	1.39	.16
Low x Males vs High x Females	4.56	1.38	2094.00	3.31	<.001

Low x Male Faces vs High x Female Faces	-1.25	1.02	2094.00	-1.23	.22
Males x Male Faces vs Females x Female Faces	-1.13	1.38	2094.00	-0.82	.41
Low x Males x Male Faces vs High x Females x Female Faces)	1.44	1.95	2094.00	0.74	.45

Random effects	**Variance**	**SD**
ID (Intercept)	316.5	17.79
Residual	105.3	10.26

Note. Baseline= Female "High good doctor" faces in females. Negative coefficients indicate lower pain intensity and unpleasantness ratings compared to the baseline condition. The random intercept captures differences in pain intensity and unpleasantness ratings in the "high good doctor" condition for female participants when the stimuli displayed are female faces.

The corresponding Bonferroni-adjusted pairwise contrasts are reported in Table S8.4.

Table S8.4

Bonferroni adjusted pairwise comparisons (EMMs) for Condition x Participant Sex x Stimulus Sex interaction on pain intensity and unpleasantness.

	Pain Intensity				
Contrasts (Sex=Male, FaceSex=Male)	**Estimate**	**SE**	**df**	**z ratio**	***p* (adjusted)**
High- Low	-3.57	1.05	2094.00	-3.40	<.01
Contrasts (Sex=Male, FaceSex=Female)					
High- Low	-2.23	1.05	2094.00	-2.12	<.05
Contrasts (Sex=Female, FaceSex=Male)					
High- Low	-0.42	0.64	2094.00	-0.66	.51
Contrasts (Sex=Female, FaceSex=Female)					
High- Low	-1.70	0.64	2094.00	-2.64	<.01

Unpleasantness					
Contrasts (Sex=Male, FaceSex=Male)	**Estimate**	**SE**	**df**	**z ratio**	**p (adjusted)**
High- Low	-1.70	0.64	2094.00	-2.64	<.01
Contrasts (Sex=Male, FaceSex=Female)					
High- Low	-2.23	1.05	2094.00	-2.12	<.05
Contrasts (Sex=Female, FaceSex=Male)					
High- Low	-0.42	0.64	2094.00	-0.66	.51
Contrasts (Sex=Female, FaceSex=Female)					
High- Low	-1.70	0.64	2094.00	-2.64	<.01

8.1.2 Main effect of Condition on pain intensity and unpleasantness

Table S8.5 provide the complete output of the linear mixed-effects model for testing the effect of Condition ("high good doctor" / "low good doctor"/ "fixation cross") on pain intensity and unpleasantness ratings.

Table S8.5

Linear- mixed effects regressions table testing the effect of Condition on pain intensity and unpleasantness.

Pain Intensity					
Fixed effects	**Estimate (β)**	**SE**	**df**	***t* value**	***p***
Intercept (Fixation cross)	50.46	1.46	146.88	34.59	<.001
High vs Fixation Cross	-0.84	0.42	3218.00	-2.02	<.05
Low vs Fixation Cross	0.72	0.42	3218.00	1.72	.08

Random effects	**Variance**	**SD**
ID (Intercept)	285.76	16.91
Residual	97.32	9.87

Unpleasantness					
Fixed effects	**Estimate (β)**	**SE**	**df**	***t* value**	***p***
Intercept (Fixation cross)	46.94	1.50	148.11	31.27	<.001
High vs Fixation Cross	-1.32	0.46	3218.00	-2.87	<.01
Low vs Fixation Cross	0.58	0.46	3218.00	1.27	.20

Random effects	**Variance**	**SD**
ID (Intercept)	300.6	17.34
Residual	118.4	10.88

N° observations =3360, ID=140

Note. Baseline=Fixation cross. Negative coefficients indicate lower pain intensity and unpleasantness ratings compared to the baseline condition. The random intercept captures individual differences in pain intensity and unpleasantness ratings in the fixation cross condition.

The corresponding Bonferroni-adjusted pairwise contrasts for all conditions are also reported in Table S8.6.

Table S8.6

Bonferroni adjusted pairwise comparisons (EMMs) for Condition on pain intensity and unpleasantness.

Pain Intensity					
Contrasts	**Estimate**	**SE**	**df**	**z ratio**	***p* (adjusted)**
Cross - High	0.84	0.42	3218	2.02	.13
Cross - Low	-0.72	0.42	3218	-1.72	.25
High - Low	-1.56	0.42	3218	-3.74	<.001

Unpleasantness					
Contrasts	**Estimate**	**SE**	**df**	**z ratio**	***p* (adjusted)**
Cross - High	1.32	0.46	3218	2.87	<.05
Cross - Low	-0.58	0.46	3218	-1.27	.61
High - Low	-1.90	0.46	3218	-4.13	<.001

8.1.3 Condition x Gene x Gene interaction: COMT rs4680 (AA vs AG vs GG) and FAAH rs324420 (A* vs CC)

Table S8.7 displays the complete output of the linear mixed-effects model for testing the effect of Condition ("high good doctor" / "low good doctor"/ "fixation cross") on pain intensity and unpleasantness ratings in individuals carrying different allelic combinations of the COMT rs4680 and FAAH rs324420 SNPs.

Table S8.7

Linear- mixed effects regressions table for pain intensity and unpleasantness in individuals carrying different combinations of the allele of COMT rs4680 and FAAH 324420.

Pain Intensity					
Fixed effects	**Estimate (β)**	**SE**	**df**	**t value**	***p***
Intercept (Fixation Cross x COMTAA x FAAH A*)	57.54	5.78	141.44	9.96	<.001
High vs Fixation Cross	-0.1	1.63	3208.00	-0.06	.95
Low vs Fixation Cross	-0.07	1.63	3208.00	-0.04	.97
COMT GA vs AA	-4.69	6.52	141.44	-0.72	.47
COMT GG vs AA	-3.44	7.96	141.44	-0.43	.67
FAAH CC vs A*	-10.17	7.31	141.44	-1.39	.17
High x COMT GA vs Fixation Cross x COMT AA	-0.68	1.84	3208.00	0.37	.71
Low x COMT GA vs Fixation Cross x COMT AA	1.68	1.84	3208.00	0.91	.36

High x COMT GG vs Fixation Cross x COMT AA	-8.54	2.25	3208.00	-3.79	<.001
Low x COMT GG vs Fixation Cross x COMT AA	-4.67	2.25	3208.00	-2.07	<.05
High x FAAH CC vs Fixation Cross x FAAH A*	-4.29	2.07	3208.00	-2.07	<.05
Low x FAAH CC vs Fixation Cross x FAAH A*	0.32	2.07	3208.00	0.16	.88
COMT (GA vs AA) x FAAH (CC vs A*)	6.43	8.37	141.44	0.77	.44
COMT (GG vs AA) x FAAH (CC vs A*)	4.12	9.62	141.44	0.43	.67
High x COMT GA x FAAH CC vs Fixation Cross x COMT AA x FAAH A*	5.84	2.37	3208.00	2.47	<.05
Low x COMT GA x FAAH CC vs Fixation Cross x COMT AA x FAAH A*	-0.21	2.37	3208.00	-.09	.92
High x COMT GG x FAAH CC vs Fixation Cross x COMT AA x FAAH A*	13.82	2.72	3208.00	5.08	<.001
Low x COMT GG x FAAH CC vs Fixation Cross x COMT AA x FAAH A*	5.11	2.72	3208.00	1.88	.06

Random effects	**Variance**	**SD**
ID (Intercept)	288.33	16.98
Residual	96.18	9.81

	Unpleasantness				
Fixed effects	**Estimate (β)**	**SE**	**df**	**t value**	***p***
Intercept (Fixation Cross x COMTAA x FAAH A*)	56.11	5.90	142.67	9.52	<.001
High vs Fixation Cross	1.06	1.79	3208.00	0.59	.56
Low vs Fixation Cross	0.01	1.79	3208.00	0.01	.99
COMT GA vs AA	-6.05	6.65	142.67	-0.91	.36
COMT GG vs AA	-7.52	8.13	142.67	-0.93	.35

FAAH CC vs A*	-14.06	7.46	142.67	-1.89	.06
High x COMT GA vs Fixation Cross x COMT AA	-1.88	2.02	3208.00	-0.93	.35
Low x COMT GA vs Fixation Cross x COMT AA	1.18	2.02	3208.00	0.58	.55
High x COMT GG vs Fixation Cross x COMT AA	-14.72	2.47	3208.00	-5.95	<.001
Low x COMT GG vs Fixation Cross x COMT AA	-4.45	2.47	3208.00	-1.80	.07
High x FAAH CC vs Fixation Cross x FAAH A*	-3.12	2.27	3208.00	-1.38	.17
Low x FAAH CC vs Fixation Cross x FAAH A*	3.99	2.27	3208.00	1.76	.08
COMT (GA vs AA) x FAAH (CC vs A*)	8.21	8.54	142.67	0.96	.34
COMT (GG vs AA) x FAAH (CC vs A*)	11.82	9.82	142.67	1.20	.23
High x COMT GA x FAAH CC vs Fixation Cross x COMT AA x FAAH A*	4.40	2.60	3208.00	1.69	.09
Low x COMT GA x FAAH CC vs Fixation Cross x COMT AA x FAAH A*	-3.87	2.60	3208.00	-1.49	.13
High x COMT GG x FAAH CC vs Fixation Cross x COMT AA x FAAH A*	16.25	2.99	3208.00	5.44	<.001
Low x COMT GG x FAAH CC vs Fixation Cross x COMT AA x FAAH A*	-0.34	2.99	3208.00	-0.12	.90

Random effects	**Variance**	**SD**
ID (Intercept)	298.40	17.27
Residual	116.00	10.77

N° observations =3360, ID=140

Note. Baseline=Fixation cross in carriers of COMT_AA e FAAH_A* alleles. Negative coefficients indicate lower pain intensity and unpleasantness ratings compared to the baseline condition. The random intercept captures individual differences in pain intensity and unpleasantness ratings in the fixation cross condition for COMT_AA /FAAH_A* carriers.

The corresponding Bonferroni-adjusted pairwise contrasts for all conditions are also reported in Table S8.8.

Table S8.8

Bonferroni adjusted pairwise comparisons (EMMs) for Condition on pain intensity and unpleasantness in individuals carrying different allelic combinations of COMT rs4680 and FAAH rs324420.

	Pain Intensity				
Contrasts (COMT=AA, FAAH=A*)	**Estimate**	**SE**	**df**	**z ratio**	**p (adjusted)**
Cross - High	0.1	1.63	∞	0.06	1.00
Cross - Low	0.07	1.63	∞	0.04	1.00
High - Low	-0.03	1.63	∞	-0.02	1.00
Contrasts (COMT=GA, FAAH=A*)					
Cross - High	0.77	0.85	∞	0.90	1.00
Cross - Low	-1.61	0.85	∞	-1.89	.18
High - Low	-2.39	0.85	∞	-2.8	<.05
Contrasts (COMT=GG, FAAH=A*)					
Cross - High	8.64	1.55	∞	5.57	<.001
Cross - Low	4.74	1.55	∞	3.06	<.01
High - Low	-3.90	1.55	∞	-2.52	<.05
Contrasts (COMT=AA, FAAH=CC)					
Cross - High	4.38	1.27	∞	3.46	<.01
Cross - Low	-0.25	1.27	∞	-0.20	1.00
High - Low	-4.63	1.27	∞	-3.66	<.001
Contrasts (COMT=AG, FAAH=CC)					
Cross - High	-0.78	0.76	∞	-1.00	.94
Cross - Low	-1.72	0.76	∞	-2.22	.08
High - Low	-0.94	0.76	∞	-1.22	.67
Contrasts (COMT=GG, FAAH=CC)					
Cross - High	-0.90	0.85	∞	-1.05	.88
Cross - Low	-0.70	0.85	∞	-0.81	1.00

High - Low	0.21	0.85	∞	0.24	1.00
	Unpleasantness				
Contrasts (COMT=AA, FAAH=A*)	**Estimate**	**SE**	**df**	**z ratio**	**p (adjusted)**
Cross - High	-1.06	1.79	∞	-0.59	1.00
Cross - Low	-0.01	1.79	∞	-0.01	1.00
High - Low	1.04	1.79	∞	-0.58	1.00
Contrasts (COMT=GA, FAAH=A*)					
Cross - High	0.82	0.94	∞	0.88	1.00
Cross - Low	-1.20	0.94	∞	-1.28	.60
High - Low	-2.02	0.94	∞	-2.15	.09
Contrasts (COMT=GG, FAAH=A*)					
Cross - High	13.66	1.70	∞	8.02	<.001
Cross - Low	4.44	1.70	∞	2.61	<.05
High - Low	-9.23	1.70	∞	-5.42	<.001
Contrasts (COMT=AA, FAAH=CC)					
Cross - High	2.07	1.39	∞	1.49	.41
Cross - Low	-4.00	1.39	∞	-2.88	<.05
High - Low	-6.07	1.39	∞	-4.36	<.001
Contrasts (COMT=AG, FAAH=CC)					
Cross - High	-0.46	0.85	∞	-0.54	1.00
Cross - Low	-1.31	0.85	∞	-1.55	.37
High - Low	-0.86	0.85	∞	-1.01	.94
Contrasts (COMT=GG, FAAH=CC)					
Cross - High	0.53	0.94	∞	0.57	1.00
Cross - Low	0.80	0.94	∞	0.85	1.00
High - Low	0.27	0.94	∞	0.28	1.00

8.1.4 Condition x Gene x Gene interaction: TPH2 rs4570625 (GG vs T*) and OXTR rs2254298 (CC vs T*)

Table S8.9 provide the complete output of the linear mixed-effects model for testing the effect of Condition (“high good doctor” / “low good doctor”/ “fixation cross”) on pain intensity and

unpleasantness ratings in individuals carrying different allelic combinations of the TPH2 rs4570625 and OXTR rs2254298 SNPs.

Table S8.9

Linear- mixed effects regressions table for pain intensity and unpleasantness in individuals carrying different combinations of the allele of TPH2 rs4570625 and OXTR rs2254298.

	Pain Intensity				
Fixed effects	**Estimate (β)**	**SE**	**df**	**t value**	**p**
Intercept (Fixation Cross x TPH2 GG x OXTR CC)	52.54	2.35	143.53	22.32	<.001
High vs Fixation Cross	-1.01	0.66	3212.00	-1.52	.13
Low vs Fixation Cross	0.55	0.66	3212.00	0.82	.41
TPH2 (T* vs GG)	-3.04	3.49	143.53	-0.87	.38
OXTR (T* vs GG)	-3.41	3.98	143.53	-0.86	.39
High x TPH2 T* vs Fixation Cross x TPH2 GG	-1.45	0.99	3212.00	-1.47	.14
Low x TPH2 T* vs Fixation Cross x TPH2 GG	-1.42	0.99	3212.00	-1.44	.15
High x OXTR T* vs Fixation Cross x OXTR CC	-0.86	1.12	3212.00	-0.76	.45
Low x OXTR T* vs Fixation Cross x OXTR CC	-1.01	1.12	3212.00	-0.90	.37
TPH2 (T* vs GG) x OXTR (T* vs CC)	1.81	6.89	143.53	0.26	.79
High x TPH2 T* x OXTR T* vs Fixation Cross x TPH2 GG x OXTR CC	11.79	1.95	3212.00	6.06	<.001
Low x TPH2 T*x OXTR T* vs Fixation Cross x TPH2 GG x OXTR CC	12.19	1.95	3212.00	6.27	<.001

Random effects	Variance	SD
ID (Intercept)	287.19	16.95
Residual	95.47	9.77

	Unpleasantness				
Fixed effects	**Estimate (β)**	**SE**	**df**	**t value**	***p***
Intercept (Fixation Cross x TPH2 GG x OXTR CC)	49.52	2.43	144.63	20.38	<.001
High vs Fixation Cross	-2.99	0.73	3212.00	-4.08	<.001
Low vs Fixation Cross	-0.24	0.73	3212.00	-0.33	0.75
TPH2 (T* vs GG)	-4.71	3.60	144.63	-1.31	.19
OXTR (T* vs GG)	-1.34	4.11	144.63	-0.33	.75
High x TPH2 T* vs Fixation Cross x TPH2 GG	1.31	1.09	3212.00	1.21	.23
Low x TPH2 T* vs Fixation Cross x TPH2 GG	0.31	1.09	3212.00	0.28	.78
High x OXTR T* vs Fixation Cross x OXTR CC	1.05	1.24	3212.00	0.84	.40
Low x OXTR T* vs Fixation Cross x OXTR CC	-1.40	1.24	3212.00	-1.13	.26
TPH2 (T* vs GG) x OXTR (T* vs CC)	-3.19	7.11	144.63	-0.45	.65
High x TPH2 T* x OXTR T* vs Fixation Cross x TPH2 GG x OXTR CC	9.71	2.14	3212.00	4.53	<.001
Low x TPH2 T*x OXTR T* vs Fixation Cross x TPH2 GG x OXTR CC	12.98	2.14	3212.00	6.05	<.001

Random effects	Variance	SD
ID (Intercept)	304.3	17.44
Residual	115.9	10.77

N° observations =3360, ID=140

Note. Baseline=Fixation cross in carriers of TPH2_GG and OXTR_CC alleles. Negative coefficients indicate lower pain intensity and unpleasantness ratings compared to the baseline condition. The random intercept captures individual differences in pain intensity and unpleasantness ratings in the fixation cross condition for TPH2_GG/OXTR_CC carriers.

The corresponding Bonferroni-adjusted pairwise contrasts for all conditions are also reported in Table S8.10.

Table S8.10

Bonferroni adjusted pairwise comparisons (EMMs) for Condition on pain intensity and unpleasantness in individuals carrying different allelic combinations of TPH2 rs4570625 and OXTR rs2254298.

	Pain Intensity				
Contrasts (TPH2=GG, OXTR=CC)	**Estimate**	**SE**	**df**	**z ratio**	**p (adjusted)**
Cross - High	1.01	0.67	∞	1.52	.39
Cross - Low	-0.55	0.67	∞	-0.82	1.00
High - Low	-1.56	0.67	∞	-2.34	.06
Contrasts (TPH2=GG, OXTR=T)					
Cross - High	1.87	0.91	∞	2.06	.12
Cross - Low	0.47	0.91	∞	0.51	1.00
High - Low	-1.40	0.91	∞	-1.54	.37
Contrasts (TPH2=T*,OXTR=CC)					
Cross - High	2.46	0.73	∞	3.38	<.01
Cross - Low	0.87	0.73	∞	1.19	.70
High - Low	-1.59	0.73	∞	-2.19	.09
Contrasts (TPH2=T*,OXTR=T*)					
Cross - High	-8.47	1.41	∞	-6.01	<.001
Cross - Low	-10.31	1.41	∞	-7.31	<001
High - Low	-1.84	1.41	∞	-1.31	.57

Unpleasantness					
Contrasts (TPH2=GG, OXTR=CC)	**Estimate**	**SE**	**df**	**z ratio**	**p (adjusted)**
Cross - High	2.99	0.73	∞	4.08	<.001
Cross - Low	0.24	0.73	∞	0.33	1.00
High - Low	-2.75	0.73	∞	-3.75	<.001
Contrasts (TPH2=GG, OXTR=T)					
Cross - High	1.94	1.00	∞	1.95	.16
Cross - Low	1.65	1.00	∞	1.65	.30
High - Low	-0.30	1.00	∞	-0.30	1.00
Contrasts (TPH2=T*,OXTR=CC)					
Cross - High	1.68	0.80	∞	2.09	.11
Cross - Low	-0.06	0.80	∞	-0.08	1.00
High - Low	-1.74	0.80	∞	-2.17	.09
Contrasts (TPH2=T*,OXTR=T*)					
Cross - High	-9.07	1.55	∞	-5.84	<.001
Cross - Low	-11.64	1.55	∞	-7.49	<.001
High - Low	-2.56	1.55	∞	-1.65	.30

8.1.5 Physiological measures: main effect of Condition on HRV components

Repeated measures ANOVA evidenced that Condition did not exert a significant effect on all the HRV components (i.e., RMSSD, VLF, LF, HF, HF/LF) considered for the experiment (RMSSD: $F(1.98, 211.84)=.87$, $p=.42$, $\eta^2_p=.01$; VLF: $F(1.76, 188.77)=1.25$, $p=.29$, $\eta^2_p=.01$; LF:$F(1.89, 202.51,)=1.75$, $p=.18$, $\eta^2_p=.02$; HF: $F(1.98, 212.17)=.12$, $p=.88$, $\eta^2_p=.001$; HF/LF:$F(1.99,213.30)=2.21$, $p=.11$, $\eta^2_p=.02$).

Post-hoc comparison results are reported in Table S8.11.

Table S8.11

Bonferroni adjusted pairwise comparison for Condition on all the HRV components considered.

	RMSSD				
Contrasts	**Mean difference**	**SE**	***p***	**Lower Bound**	**Upper Bound**
High - Low	2.46	1.94	.63	-2.27	7.19
- Cross	0.49	2.07	1.00	-4.55	5.52
Low - High	-2.46	1.94	.63	-7.09	2.27
- Cross	-1.97	1.91	.91	-6.61	2.67
Cross - High	-0.49	2.07	1.00	-5.52	4.55
- Low	1.97	1.91	.91	-2.67	6.61
	VLF				
Contrasts	**Mean difference**	**SE**	***p***	**Lower Bound**	**Upper Bound**
High - Low	25.04	22.16	.78	-28.84	78.93
- Cross	-6.75	23.85	1.00	-64.76	51.26
Low - High	-25.04	22.16	.78	-78.93	28.84
- Cross	-31.79	17.09	.19	-73.37	9.78
Cross - High	6.74	23.85	1.00	-51.26	64.76
- Low	31.79	17.09	.20	-9.78	73.37
	LF				
Contrasts	**Mean difference**	**SE**	***p***	**Lower Bound**	**Upper Bound**
High - Low	199.08	108.33	.21	-64.39	462.56
- Cross	62.97	119.70	1.00	-228.16	354.11
Low - High	-199.08	108.33	.21	-462.56	64.39
- Cross	-136.11	97.20	.49	-372.52	100.29
Cross - High	-62.97	119.700	1.00	-354.11	228.16
- Low	136.11	97.20	.49	-100.29	372.52

Contrasts	HF Mean difference	SE	*p*	Lower Bound	Upper Bound
High - Low	26.54	88.32	1.00	-188.28	241.90
- Cross	-17.69	94.26	1.00	-246.95	211.57
Low - High	-26.54	88.32	1.00	-241.36	188.27
- Cross	-44.23	87.81	1.00	-257.79	169.33
Cross - High	17.69	94.26	1.00	-211.57	246.95
- Low	44.23	87.81	1.00	-169.33	257.79
Contrasts	**LF/HF Mean difference**	**SE**	***p***	**Lower Bound**	**Upper Bound**
High - Low	0.23	0.12	.16	-0.05	0.50
- Cross	-0.05	0.12	1.00	-0.24	0.33
Low - High	-0.22	0.12	.16	-0.51	0.05
- Cross	-0.18	0.11	.32	-0.45	0.09
Cross - High	0.05	0.12	1.00	-0.03	0.24
- Low	0.18	0.1	.32	-0.08	0.45

8.1.6 Questionnaires: correlational analyses between questionnaires scorings

Table S8.12 reports the correlations between the different questionnaires scorings.

Table S8.12

Correlations between questionnaires scorings.

	Positive Affect Pre pain	Negative Affect Pre pain	Positive Affect Post pain	Negative affect Post pain
PANAS				
Positive Affect pre pain	1	-.10	.70***	-.03
Negative Affect pre pain	-.10	1	-.16	.70***
Positive Affect post pain	.70***	-.16	1	-.16
Negative Affect post pain	-.03	.70***	-.16	1

STAI				
Trait Anxiety	-.26	.40***	-.36**	.29***
State Anxiety	-.30***	.68	.10***	.55***
FPQ				
Fear of Pain	.07	.13	-.22	.22
PCS				
Pain Catastrophizing	-.16	.22**	-.07	.18
BFI				
Agreeableness	.06	-.02	-.01	-.11
Conscientiousness	.31***	-.12	.30	-.06
Emotional Stability	.14	-.40***	.14	-.31***
Extraversion	.15	-.22	.24**	-.25**
Openness	.20*	.04	.22	.10
BIS/BAS				
Behavioural Inhibition	.002	.23	-.12	.26
Drive	.21*	-.05	.20*	-.03
Fun	.04	.02	.11	.05
Reward	.26**	-.04	.26	-.09
Importance trust	.20*	.03	.11	.12

	Trait Anxiety	**State Anxiety**	**Fear of Pain**	**Pain Catastrophizing**
PANAS				
Positive Affect pre pain	-.26**	-.30***	.07	-.16
Negative Affect pre pain	.40*	.68***	13	.22
Positive Affect post pain	-.22	-.36***	.10	-.07
Negative Affect post pain	.29***	.55***	.22*	.18*
STAI				
Trait Anxiety	1	.49***	.07	.35***
State Anxiety	.49	1	.09	.17

FPQ				
Fear of Pain	.07	.09	1	.33***
PCS				
Pain Catastrophizing	.35***	.17	.33***	1
BFI				
Agreeableness	-.31***	-.13	-.01	-.10
Conscientiousness	-.30***	-.19*	.02	-.14
Emotional Stability	-.65*	-.45***	-.22	-.36***
Extraversion	-.44***	-.29***	.08	-.05
Openness	.08	.03	-.01	.03
BIS/BAS				
Behavioural Inhibition	.63***	.41***	.28***	.37***
Drive	-.21*	-.10	.21	.08
Fun	.15	.08	.10	.08
Reward	.005	-.02	.13	.10
Importance trust	-.24	-.08	.11	-.21*

	Agreeableness	Conscientiousness	Emotional Stability	Extraversion
PANAS				
Positive Affect pre pain	.06	.31***	.14	.15
Negative Affect pre pain	-.02	-.12	.40***	-.22
Positive Affect post pain	-.005	.30***	.14	.24
Negative Affect post pain	-.11	-.06	-.31***	-.25
STAI				
Trait Anxiety	-.31***	-.30***	-.65***	-.29***
State Anxiety	-.13	-.19*	-.45***	-.44***
FPQ				
Fear of Pain	-.01	.02	-.22	.08
PCS				
Pain Catastrophizing	-.10	-.14	-.36***	-.05
BFI				
Agreeableness	1	.09	.20*	.25***
Conscientiousness	.09	1	.21	.26**
Emotional Stability	.20*	.21	1	.28***
Extraversion	.25***	.26***	.28***	1
Openness	-.03	.05	-.06	-.09
BIS/BAS				
Behavioural Inhibition	-.09	-.06	-.56***	-.25**
Drive	-.13	.20*	.03	.07
Fun	.05	-.14	-.13	.04
Reward	.08	.19*	-.10	.17*
Importance trust	.19*	.15	.14	.25***

	Openness	Behavioural Inhibition	Drive	Fun
PANAS				
Positive Affect pre pain	.20	.002	.21	.04
Negative Affect pre pain	.04	.23*	-.05	.02
Positive Affect post pain	.22*	-.12	.20	.11
Negative Affect post pain	.10	.26**	-.03	-.05
STAI				
Trait Anxiety	.08	.63***	-.10	.15
State Anxiety	.03	.41***	-.21*	.08
FPQ				
Fear of Pain	-.01	.28***	.21	.10
PCS				
Pain Catastrophizing	.03	.37***	.08	.08
BFI				
Agreeableness	-.03	-.09	-.13	.05
Conscientiousness	.05	-.06	.20	-.14
Emotional Stability	-.06	-.56***	.03	-.13
Extraversion	-.09	-.25**	.07	.04
Openness	1	.03	-.04	.15
Behavioural Inhibition			-.05	.10
BIS/BAS	.03	1		
Drive	-.04	-.05	1	.18*
Fun	.15	.10	.18	1
Reward	.15	.24**	.12	.31*
Importance trust	-.01	-.06	.17	.06

	Reward	Importance trust
PANAS		
Positive Affect pre pain	.26**	.20*
Negative Affect pre pain	-.04	.03
Positive Affect post pain	.26**	.11
Negative Affect post pain	-.09	.12
STAI		
Trait Anxiety	.01	-.08
State Anxiety	-.02	-.24**
FPQ		
Fear of Pain	.13	.12
PCS		
Pain Catastrophizing	.10	-.21
BFI		
Agreeableness	.08	.19*
Conscientiousness	.19*	.15
Emotional Stability	-.10	.14
Extraversion	.17*	.25**
Openness	.15	-.01
Behavioural Inhibition	.24**	-.06
BIS/BAS		
Drive	.12	.17
Fun	.31***	.06
Reward	1	.24**
Importance trust	.24**	1

Note. *p<.05, ** p<.01, *** p<.001

8.2 Chapter 5: Additional results

8.2.1 Idiosyncratic "good doctor" models: Exploratory analyses on prevalent sex and age of the idiosyncratic "good doctors"

Although not directly related to our main hypothesis, we performed exploratory analyses to further examine potential trends in the data. These analyses were not based on preregistered hypotheses and must therefore be interpreted with caution. Their purpose is to provide a more comprehensive picture of the data, highlighting possible patterns which may inform future research on the impact of first impressions on clinical outcomes.

Previous findings suggest that female physicians are perceived as more patient- centered, but less competent than their male counterparts (Hall et al., 2020), with this difference becoming more pronounced when older vs. younger male physicians are considered. In respect of age, data are more mixed: while some studies report that older physicians are perceived as warm but incompetent (Cuddy & Fiske, 2004; Fiske et al., 2002), more recent findings suggest the opposite (Hall et al., 2020). In light of this literature, our exploratory analyses propose to observe whether a prevalence in terms of sex and age emerged within participants' idiosyncratic models of the "good doctor" stereotype developed in Experiment 1.

8.2.1.1. Data analyses

Perceived age and sex of the faces generated using the idiosyncratic models were estimated using a pretrained classifier adapted from Peterson et al. (2022;Fujita et al., 2025), which is capable of predicting various physical and social attributes using the latent features associated with each face.

Notably, these analyses were performed considering only images representing the linear interpolation at +1 SD (i.e., "high good doctor" faces) and -1 SD (i.e., "low good doctor" faces), without taking into account images at intermediate levels of the model (i.e., faces representing the linear interpolation at ± 0.5 from the mean) which reflected partial manipulations.

To examine whether there were significant differences in the frequency with which male versus female "high good doctor" images were generated, as well as the influence of participant biological sex on the tendency to generate male or female "high good doctor" and "low good doctor" faces, one-sample proportion tests and Chi-square tests were conducted. Additionally, to investigate any

differences in the perceived average age between "high good doctor" and "low good doctor" faces, an independent-samples *t*-test was performed.

The significance level was set at $p<0.05$ for all analyses.

Data analysis was conducted using RStudio.

8.2.1.2 Results

A one-sample proportion test showed that the proportions of generated "high good doctor" male ($\hat{p}_m$=0.71) and female faces ($\hat{p}_{♀}$ =0.29) differed significantly from the expected value (p_0=0.50) (χ^2=15.21, *df*=1, $p<.001$), with "high good doctor" male faces generated more frequently than female ones. This pattern held across participant biological sex, as no significant differences were found in the frequency of "high good doctor" male and female faces developed from male and female participants (χ^2=1.59, *df*=1, *p*=.21).

For "low good doctor" faces, no significant differences emerged between the observed and expected proportions (p_0=0.50) of male ($\hat{p}_m$=0.54) and female ($\hat{p}_{♀}$=0.46) faces generated (χ^2=0.28, *df*=1, *p*=.60). Likewise, male and female participants did not differ in their tendency to generate male or female "low good doctor" faces (χ^2=0.25, df=1, *p*=.61).

Regarding perceived age, an independent sample T-Test evidenced a significant difference in the average $Zscore_{_age}$ values between "high good doctor" (M_{age}=0.47) and "low good doctor" faces (M_{age}=-0.51) (*t*=9.86, *df*=177.98, $p<.001$), with the former being associated with higher perceived age on average.

8.2.2 Idiosyncratic "good doctor" models: Exploratory analyses on the social attributes underlying the idiosyncratic "good doctors"

Exploratory analyses were conducted to explore the social traits underlying participants' idiosyncratic representations of the "good doctor". Specifically, we wanted to observe whether these individual-level, stereotypical representations would reflect the same attributes that emerged as defining the group-level "good doctor" dimension identified in our previous study.

8.2.2.1 Data analysis

Analyses were performed considering the stimuli administered during the second phase of the study and their corresponding ratings. Stimuli generated using the models of participants who were excluded from the second phase were also removed, to ensure that each included stimulus had been evaluated both by the participant whose ratings were used to generate its model and by five other participants.

For each stimulus, we computed the mean of the ratings participants assigned depending on their resemblance with a "good doctor". These mean ratings were then transformed in Z- scores. Correlations were subsequently computed between these Z- scores and the Z-scores of the same images across the 36 attributes. Notably, the Z-scores for each of the 36 personality traits were estimated using StyleGAN2 generative models.

8.2.2.2 Results

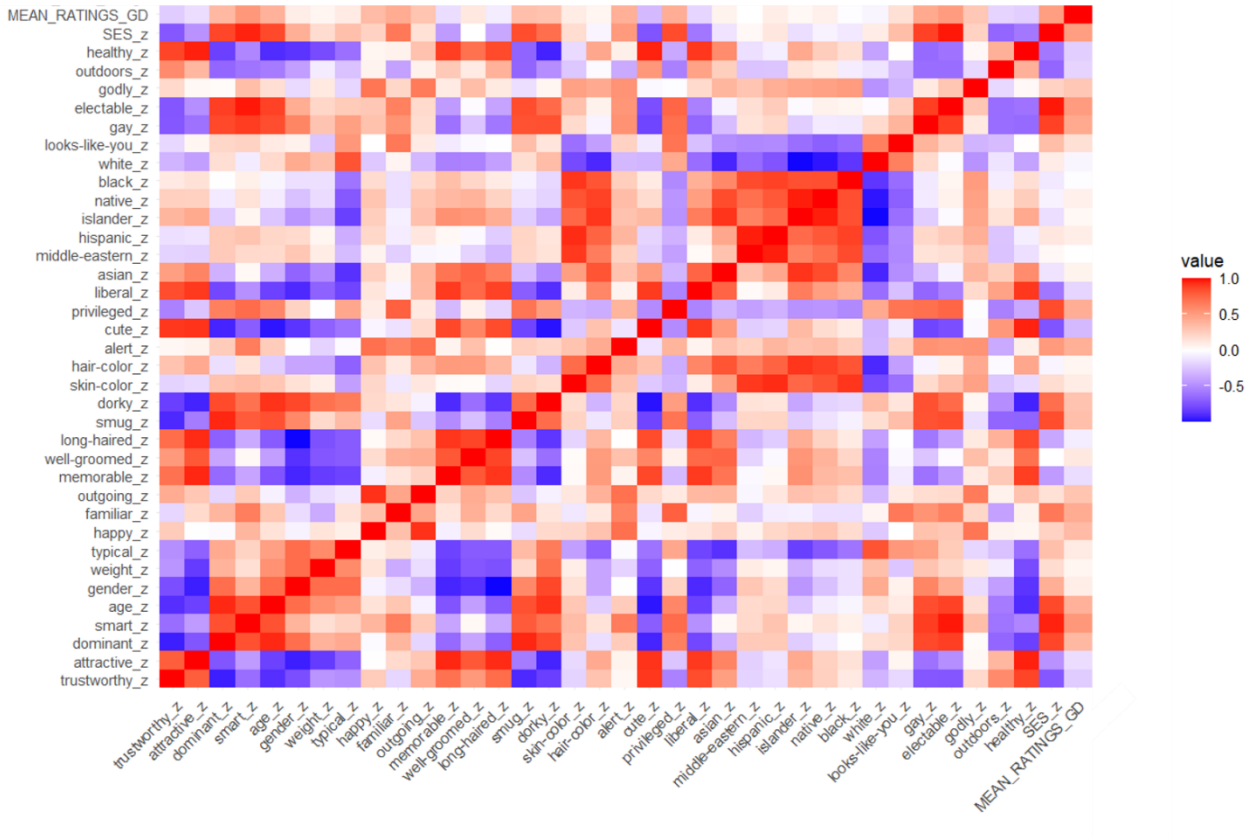
Results partially replicate the correlations we have observed in our previous study.

We found that participants' ratings assigned to each facial image based on its perceived resemblance with a "good doctor" were positively correlated with the "familiar" (r=.43), "smart" (r=.53), "alert" (r=.41), "privileged" (r=.42) and "electable" (r=.51) attributes, corroborating previous results. In addition, positive correlations emerged with the "socio-economic status" (r=.50), "age" (r=.40) and "gay" (r=.44) attributes. Weaker correlations also emerged with the "dominant" (r=.36), "happy" (r=.35), "smug" (r=.33), "dorky" (r=.31), "cute" (r=.30), "outgoing" (r=.30) and "godly" (r=.29) social traits.

All correlations are displayed in Figure S8.2.

Figure S8.2

Correlations between the images ratings according to their resemblance with a "good doctor" and the 34 attributes of the One Million Impression Dataset (Peterson et al., 2022).



8.3 Chapter 6: Additional results

8.3.1 Main effect of Condition on pain intensity and unpleasantness

Table S8.13 provide the complete output of the linear mixed-effects model for testing the effect of Condition (idiosyncratic "high good doctor" / idiosyncratic "low good doctor"/ other participant's "high good doctor"/ other participant's "low good doctor"/ "fixation cross") on pain intensity and unpleasantness ratings.

Table S8.13

Linear- mixed effects regressions table testing the effect of Condition on pain intensity and unpleasantness.

	Pain Intensity				
Fixed effects	**Estimate (β)**	**SE**	**df**	***t* value**	***p***
Intercept (Fixation cross)	58.21	2.93	31.63	19.85	<.001
High idiosyncratic vs Fixation Cross	-1.32	0.96	866.00	-1.39	.17
High other vs Fixation Cross	0.19	0.96	866.00	0.20	.84
Low idiosyncratic vs Fixation Cross	0.19	0.96	866.00	0.20	.84
Low other vs Fixation Cross	-1.59	0.96	866.00	-1.67	.10

Random effects	**Variance**	**SD**
ID (Intercept)	244.15	15.63
Residual	82.28	9.07

	Unpleasantness				
Fixed effects	**Estimate (β)**	**SE**	**df**	***t* value**	***p***
Intercept (High idiosyncratic vs Fixation Cross)	48.69	3.27	30.76	14.85	<.001
High other vs Fixation Cross	-3.64	0.88	866.00	-4.13	<.001
Low idiosyncratic vs Fixation Cross	-2.75	0.88	866.00	-3.12	<.01
Low other vs Fixation Cross	-1.10	0.88	866.00	-1.25	.21
High idiosyncratic vs Fixation Cross	-3.21	0.88	866.00	-3.63	<.001

Random effects	**Variance**	**SD**
ID (Intercept)	310.77	17.69
Residual	70.09	8.37

N°observations=900, ID=30

Note. Baseline=Fixation cross. Negative coefficients indicate lower pain intensity and unpleasantness ratings compared to the baseline condition. The random intercept captures individual differences in pain intensity and unpleasantness ratings in the fixation cross condition.

The corresponding Bonferroni-adjusted pairwise contrasts for all conditions are also reported in Table S8.14.

Table S8.14

Bonferroni adjusted pairwise comparisons (EMMs) for Condition on pain intensity and unpleasantness.

Pain Intensity					
Contrasts	**Estimate**	**SE**	**df**	**z ratio**	***p* (adjusted)**
Cross - High idiosyncratic	1.33	0.956	866	1.39	1.00
Cross - High other	-0.19	0.956	866	-0.20	1.00
Cross - Low idiosyncratic	-0.19	0.956	866	-0.20	1.00
Cross - Low other	1.59	0.956	866	1.67	.96
High idiosyncratic – High other	-1.52	0.956	866	-1.59	1.00
High idiosyncratic – Low idiosyncratic	-1.52	0.956	866	-1.59	1.00
High idiosyncratic – Low other	0.26	0.956	866	0.28	1.00
High other – Low idiosyncratic	0.01	0.956	866	0.001	1.00
High other – Low other	1.79	0.956	866	1.87	.62
Low idiosyncratic- Low other	1.78	0.956	866	1.87	.63

Unpleasantness					
Contrasts	**Estimate**	**SE**	**df**	**z ratio**	***p* (adjusted)**
Cross - High idiosyncratic	3.64	0.88	866	4.13	.000
Cross - High other	2.75	0.88	866	3.12	.02
Cross - Low idiosyncratic	1.11	0.88	866	1.25	1.00
Cross - Low other	3.21	0.88	866	3.64	.003

High idiosyncratic – High other	-0.89	0.88	866	-1.01	1.00
High idiosyncratic – Low idiosyncratic	-2.54	0.88	866	-2.88	.04
High idiosyncratic – Low other	-0.43	0.88	866	-0.49	1.00
High other – Low idiosyncratic	-1.64	0.88	866	-1.86	.63
High other – Low other	0.46	0.88	866	0.52	1.00
Low idiosyncratic- Low other	2.11	0.88	866	2.38	.17

Note. High idio= idiosyncratic "high good doctor" condition; High other= other participant's "high good doctor" condition; Low idio= idiosyncratic "low good doctor" condition; Low other= other participant's "low good doctor" condition

8.3.2 Pain Task: Main effect of Condition on Mean HR and HRV components (i.e., RMSSD,VLF,LF,HF,LF/HF)

Table S8.15 shows the Bonferroni- adjusted pairwise comparisons for all the physiological measures considered in the study (i.e., Mean HR, RMSSD, VLF, LF, HF, HF/LF).

Table S8.15

Bonferroni- adjusted pairwise comparisons for all the physiological measures considered in the study (i.e., Mean HR, RMSSD, VLF, LF, HF, HF/LF)

	MeanHR				
Contrasts	**Mean difference**	**SE**	***p***	**Lower Bound**	**Upper Bound**
High idiosyncratic - Low idiosyncratic	0.78	0.72	1.00	-1.51	3.08
- Cross	1.20	0.52	.33	-0.45	2.84
- High other	0.79	0.66	1.00	-1.30	2.87
- Low other	0.63	0.61	1.00	-1.29	2.55
Low idiosyncratic - High idiosyncratic	-0.79	0.72	1.00	-3.08	1.51
- Cross	0.41	0.53	1.00	-1.25	2.07
- High other	0.001	0.46	1.00	-1.44	1.44
- Low other	-0.16	0.50	1.00	-1.72	1.41
Cross - High idiosyncratic	-1.19	0.52	.33	-2.84	0.45

- Low idiosyncratic	-0.41	0.53	1.00	-2.07	1.25
- High other	-0.41	0.34	1.00	-1.49	0.67
- Low other	-0.57	0.31	.85	-1.55	0.42
High other - High idiosyncratic	-0.79	0.66	1.00	-2.87	1.30
- Low idiosyncratic	-0.001	0.46	1.00	-1.44	1.44
- Cross	0.41	0.34	1.00	-0.67	1.49
- High other	-0.16	0.46	1.00	-1.60	1.28
Low other - High idiosyncratic	-0.63	0.61	1.00	-2.55	1.28
- Low idiosyncratic	0.16	0.50	1.00	-1.41	1.72
- Cross	0.57	0.31	.85	-0.42	1.55
- High other	0.16	0.46	1.00	-1.28	1.60

Contrasts	**RMSSD Mean difference**	**SE**	***p***	**Lower Bound**	**Upper Bound**
High idiosyncratic - Low idiosyncratic	-0.24	1.01	1.00	-3.41	2.93
- Cross	-0.52	1.57	1.00	-5.48	4.45
- High other	0.08	1.73	1.00	-5.54	5.38
- Low other	-0.97	1.71	1.00	-6.36	4.41
Low idiosyncratic - High idiosyncratic	0.24	1.01	1.00	-2.93	3.41
- Cross	-0.28	1.41	1.00	-4.73	4.18
- High other	0.16	1.52	1.00	-4.64	4.96
- Low other	-0.73	1.49	1.00	-5.43	3.97
Cross - High idiosyncratic	0.52	1.57	1.00	-4.45	5.47
- Low idiosyncratic	0.28	1.41	1.00	-4.18	4.73
- High other	0.44	1.25	1.00	-3.51	4.38
- Low other	-0.46	2.63	1.00	-8.77	7.86
High other - High idiosyncratic	-0.08	1.73	1.00	-5.38	5.54
- Low idiosyncratic	-0.16	1.52	1.00	-4.96	4.64
- Cross	-0.44	1.25	1.00	-4.38	3.51
- High other	-0.89	2.53	1.00	-8.88	7.09
Low other - High idiosyncratic	0.97	1.71	1.00	-4.42	6.36
- Low idiosyncratic	0.73	1.49	1.00	-3.97	5.43

- Cross	0.46	2.64	1.00	-7.86	8.77
- High other	0.89	2.53	1.00	-7.09	8.88

Contrasts	**VLF Mean difference**	**SE**	***p***	**Lower Bound**	**Upper Bound**
High idiosyncratic - Low idiosyncratic	2.87	29.85	1.00	-91.26	97.01
- Cross	-0.68	19.33	1.00	-61.66	60.30
- High other	-7.55	20.93	1.00	-73.54	58.44
- Low other	11.79	19.04	1.00	-48.24	71.83
Low idiosyncratic - High idiosyncratic	-2.87	29.85	1.00	-97.01	91.26
- Cross	-3.55	20.78	1.00	-69.09	61.99
- High other	-10.42	24.24	1.00	-86.87	66.02
- Low other	8.92	23.89	1.00	-66.40	84.24
Cross - High idiosyncratic	0.68	19.33	1.00	-60.30	61.66
- Low idiosyncratic	3.54	20.78	1.00	-61.98	69.09
- High other	-6.87	16.10	1.00	-57.63	43.88
- Low other	12.47	15.70	1.00	-37.02	61.96
High other - High idiosyncratic	7.55	20.92	1.00	-58.44	73.54
- Low idiosyncratic	10.42	24.24	1.00	-66.02	86.87
- Cross	6.87	16.10	1.00	-43.88	57.63
- High other	19.34	18.04	1.00	-37.54	76.22
Low other - High idiosyncratic	-11.79	19.04	1.00	-71.83	48.25
- Low idiosyncratic	-8.92	23.89	1.00	-84.24	66.40
- Cross	-12.47	15.70	1.00	-61.96	37.02
- High other	-19.34	18.04	1.00	-76.22	37.54

Contrasts	**LF Mean difference**	**SE**	***p***	**Lower Bound**
High idiosyncratic - Low idiosyncratic	-26.26	108.88	1.00	-369.60
- Cross	70.23	133.73	1.00	-351.47
- High other	-177.78	170.38	1.00	-715.04
- Low other	119.59	158.44	1.00	-380.04

Low idiosyncratic - High idiosyncratic	26.26	108.88	1.00	-317.06	
- Cross	96.49	154.13	1.00	-389.54	
- High other	-151.52	121.99	1.00	-536.20	
- Low other	145.85	130.86	1.00	-266.79	
Cross - High idiosyncratic	-70.23	133.73	1.00	-491.92	
- Low idiosyncratic	-96.49	154.13	1.00	-582.51	
- High other	-248.00	197.17	1.00	-869.76	
- Low other	49.36	179.19	1.00	-515.70	
High other - High idiosyncratic	177.78	170.38	1.00	-359.48	
- Low idiosyncratic	151.52	121.99	1.00	-233.17	
- Cross	248.00	197.17	1.00	-373.75	
- High other	297.37	196.77	1.00	-323.12	
Low other - High idiosyncratic	-119.59	158.44	1.00	-619.22	
- Low idiosyncratic	-151.52	121.99	1.00	-233.17	
- Cross	-49.36	179.19	1.00	-614.63	
- High other	-297.37	196.77	1.00	-917.85	

Contrasts	**HF Mean difference**	**SE**	***p***	**Lower Bound**	**Upper Bound**
High idiosyncratic - Low idiosyncratic	29.05	36.72	1.00	-86.75	144.85
- Cross	-55.58	58.00	1.00	-238.48	127.32
- High other	6.53	67.22	1.00	-205.46	218.52
- Low other	-20.29	54.64	1.00	-192.61	152.02
Low idiosyncratic - High idiosyncratic	-29.05	36.72	1.00	-144.85	86.75
- Cross	-84.64	50.37	1.00	-243.46	74.19
- High other	-22.52	56.32	1.00	-200.13	155.84
- Low other	-49.35	47.90	1.00	-200.39	101.69
Cross - High idiosyncratic	55.58	58.00	1.00	-127.32	238.49
- Low idiosyncratic	84.64	50.36	1.00	-74.19	243.46
- High other	62.11	48.15	1.00	-89.72	213.94
- Low other	35.29	87.30	1.00	-240.01	310.58
High other - High idiosyncratic	-6.53	67.22	1.00	-218.51	205.46

- Low idiosyncratic	22.53	56.32	1.00	-155.84	200.13
- Cross	-62.11	48.15	1.00	-213.94	89.72
- High other	-26.83	86.49	1.00	-299.57	245.92
Low other - High idiosyncratic	20.29	54.64	1.00	-152.02	192.61
- Low idiosyncratic	49.35	47.89	1.00	-101.70	200.39
- Cross	-35.29	87.30	1.00	-310.58	240.01
- High other	26.82	86.49	1.00	-245.93	299.57

Contrasts	**LF/HF Mean difference**	**SE**	***p***	**Lower Bound**	**Upper Bound**
High idiosyncratic - Low idiosyncratic	-0.04	0.33	1.00	-1.09	1.02
- Cross	0.61	0.24	.20	-0.15	1.36
- High other	-0.15	0.31	1.00	-1.14	0.83
- Low other	0.41	0.36	1.00	-0.72	1.55
Low idiosyncratic - High idiosyncratic	0.04	0.33	1.00	-1.02	1.09
- Cross	0.64	0.34	.73	-0.43	1.71
- High other	-0.12	0.30	1.00	-1.06	0.83
- Low other	0.45	0.18	.24	-0.13	1.03
Cross - High idiosyncratic	-0.61	0.24	.20	-1.36	0.15
- Low idiosyncratic	-0.64	0.34	.73	-1.71	0.42
- High other	-0.76	0.32	.24	-1.78	0.26
- Low other	-0.19	0.28	1.00	-1.06	0.68
High other - High idiosyncratic	0.15	0.31	1.00	-0.83	1.14
- Low idiosyncratic	0.12	0.30	1.00	-0.83	1.06
- Cross	0.76	0.32	.30	-0.26	1.78
- High other	0.57	0.34	1.00	-0.49	1.63
Low other - High idiosyncratic	-0.41	0.36	1.00	-1.55	0.72
- Low idiosyncratic	-0.45	0.18	.24	-1.03	0.13
- Cross	0.19	0.28	1.00	-0.68	1.06
- High other	-0.57	0.34	1.00	-1.63	0.49

Note. High idio= idiosyncratic "high good doctor" condition; High other= other participant's "high good doctor" condition; Low idio= idiosyncratic "low good doctor" condition; Low other= other participant's "low good doctor" condition

8.3.3 Familiarity task: Main effect of Condition on Familiarity ratings

Table S8.16 illustrates the Bonferroni- adjusted pairwise comparisons for the familiarity ratings across all conditions considered (idiosyncratic "high good doctor"/ "low good doctor"; other participant's "high good doctor" / "low good doctor"; cross).

Table S8.16

Bonferroni- adjusted pairwise comparisons for the familiarity ratings across all conditions

FAMILIARITY RATINGS					
Contrasts	**Mean difference**	**SE**	***p***	**Lower Bound**	**Upper Bound**
High idiosyncratic - Low idiosyncratic	-1.60	6.00	1.00	-22.22	19.02
- High other	-1.93	5.76	1.00	-21.75	17.88
- Low other	-3.63	4.56	1.00	-19.31	12.04
- High other similar	10.33	6.66	1.00	-12.59	33.26
- Low other similar	20.83	6.23	.06	-0.60	42.26
- High other different	19.80	7.37	.33	-5.54	45.14
- Low other different	18.27	7.41	.56	-7.24	43.77
Low idiosyncratic - High idiosyncratic	1.6	6.00	1.00	-19.02	22.22
- High other	-0.33	6.14	1.00	-21.45	20.78
- Low other	-2.03	4.96	1.00	-19.10	15.04
- High other similar	11.93	7.99	1.00	-15.54	39.41
- Low other similar	22.43	6.46	.05	0.23	44.64
- High other different	21.40	8.18	.39	-6.73	49.53
- Low other different	19.87	8.14	.59	-8.12	47.85
High other- High idiosyncratic	1.93	5.76	1.00	-17.88	21.75
- Low idiosyncratic	0.33	6.14	1.00	-20.78	21.45
- Low other	-1.70	5.81	1.00	-21.70	18.30

- High other similar	12.27	7.50	1.00	-13.55	38.08
- Low other similar	22.77	5.88	.02	2.55	42.99
- High other different	21.73	6.59	.07	-0.93	44.40
- Low other different	20.20	6.72	.15	-2.93	43.40
Low other - High idiosyncratic	3.63	4.56	1.00	-12.04	19.31
- Low idiosyncratic	2.03	4.96	1.00	-15.04	19.10
- Low other	1.70	5.81	1.00	-18.30	21.70
- High other similar	13.97	6.31	.98	-7.73	35.66
- Low other similar	24.47	5.40	.003	5.89	43.04
- High other different	23.43	7.61	.13	-2.72	49.59
- Low other different	21.90	7.12	.13	-2.58	46.38
Low other - High idiosyncratic	3.63	4.56	1.00	-12.04	19.31
- Low idiosyncratic	2.03	4.96	1.00	-15.04	19.10
- High other	1.70	5.81	1.00	-18.30	21.70
- High other similar	13.97	6.31	.98	-7.73	35.66
- Low other similar	24.47	5.40	.003	5.89	43.04
- High other different	23.43	7.61	.13	-2.72	49.59
- Low other different	21.90	7.12	.13	-2.58	46.38
High other similar - High idiosyncratic	-10.33	6.66	1.00	-33.26	12.59
- Low idiosyncratic	-11.93	7.99	1.00	-39.40	15.54
- High other	-12.27	7.50	1.00	-38.08	13.55
- Low other	-13.97	6.31	.98	-35.66	7.73
- Low other similar	10.50	5.48	1.00	-8.34	29.34
- High other different	9.47	6.33	1.00	-12.31	31.24
- Low other different	7.93	5.73	1.00	-11.78	27.65
Low other similar - High idiosyncratic	-20.83	6.23	.06	-42.26	0.60
- Low idiosyncratic	-22.43	6.46	.05	-44.64	-0.23
- High other	-22.77	5.88	.02	-42.99	-2.55
- Low other	-24.47	5.40	.003	-43.04	-5.89
- High other similar	-10.50	5.48	1.00	-29.34	8.34
- High other different	-1.03	5.41	1.00	-19.64	17.58
- Low other different	-2.57	5.17	1.00	-20.34	15.21

High other different - High idiosyncratic	-19.80	7.37	.33	-45.14	5.54
- Low idiosyncratic	-21.40	8.18	.39	-49.53	6.73
- High other	-21.73	6.59	.07	-44.40	0.93
- Low other	-23.43	7.61	.13	-49.59	2.72
- High other similar	-9.47	6.33	1.00	-31.24	12.31
- Low other similar	1.03	5.41	1.00	-17.58	19.64
- Low other different	-1.53	4.40	1.00	-16.68	13.61
Low other different - High idiosyncratic	-18.27	7.41	.56	-43.77	7.24
- Low idiosyncratic	-19.87	8.14	.59	-47.85	8.12
- High other	-20.20	6.72	.15	-43.33	2.93
- Low other	-21.90	7.12	.12	-46.38	2.58
- High other similar	-7.93	5.73	1.00	-27.65	11.78
- Low other similar	2.57	5.17	1.00	-15.21	20.34
- High other different	1.53	4.40	1.00	-13.61	16.68

Note. High/Low other_sim= “high” and “low good doctor” faces generated using the model of another participant which resulted to be the most similar; High/low other_diff= “high” and “low good doctor” faces generate using the model of another participant which resulted to be very different.

8.3.4 Questionnaires: correlational analyses between pain measures and questionnaires scorings

Table S8.17 reports the correlations between pain measures (i.e., pain intensity and unpleasantness across conditions (“high good doctor”/ “low good doctor” / “fixation cross”)) and questionnaires scorings.

Table S8.17

Correlations between questionnaires scorings.

	Positive Affect pre pain	Negative Affect pre pain	Positive Affect post pain	Negative affect post pain
PANAS				
Positive Affect pre pain	1	.02	.79***	.18
Negative Affect pre pain	.02	1	-.06	.61***
Positive Affect post pain	.79***	-.06	1	.17
Negative affect post pain	.18	.61***	.17	1
STAI				
Trait Anxiety	-.38*	.39	-.33	.40*
State Anxiety	-.48**	.38	-.49**	.14
FPQ				
Fear of Pain	-.09	-.39	-.13	-.09
PCS				
Pain Catastrophizing	-.01	.18	-.14	.28
BFI				
Agreeableness	.05	-.18	.17	-.30
Conscientiousness	.31	-.31	.32	-.16
Emotional Stability	-.08	.39*	-.24	.29
Extraversion	.37*	.39*	.06	.16
Openness	.46	-.08	.42*	-.06
BIS/BAS				
Behavioural Inhibition	.22	-.08	.43*	.01
Drive	-.27	-.34	-.18	-.53**
Fun	-.26	-.42*	-.20	-.25
Reward	-.13	-.38*	-.08	-.37*

Importance trust	.11	.23	-.14	.16

	Trait Anxiety	State Anxiety	Fear of Pain	Pain Catastrophizing
PANAS				
Positive Affect pre pain	-.49**	-.37*	-.09	.42
Negative Affect pre pain	.39*	.38*	-.39*	.96
Positive Affect post pain	-.49*	-.33	-.13	.35
Negative affect post pain	.14	.40*	-.09	.48
STAI				
Trait Anxiety	1	.61***	.04	.14
State Anxiety	.61***	1	.14	.09
FPQ				
Fear of Pain	.04	.14	1	.14
PCS				
Pain Catastrophizing	.32	.28	.52**	.004
BFI				
Agreeableness	-.34	-.30	-.05	-.26
Conscientiousness	-.48**	-.24	-.17	-.36
Emotional Stability	.71***	.58***	.11	.38
Extraversion	-.20	-.05	-.19	.13
Openness	-.06	-.12	-.13	.11
BIS/BAS				
Behavioural Inhibition	-.64***	-.43	-.42	-.39*
Drive	-.13	-.27	-.004	-.26
Fun	-.02	-.14	.02	-.30
Reward	-.19	-.32	-.01	-.22
Importance trust	-.16	-.03	.14	.20

	Agreeableness	Conscientiousness	Emotional Stability	Extraversion
PANAS				
Positive Affect pre pain	.05	.31	-.08	.37
Negative Affect pre pain	-.18	-.31	.39*	.39
Positive Affect post pain	.17	.32	-.24	.06
Negative affect post pain	-.30	-.16	.29	.16
STAI				
Trait Anxiety	-.34	-.48	.71***	-.20
State Anxiety	-.30	-.24	.58***	-.05
FPQ				
Fear of Pain	-.05	-.17	.11	-.19
PCS				
Pain Catastrophizing	-.26	-.36	.38	.13
BFI				
Agreeableness	1	.02	-.45	-.02
Conscientiousness	.02	1	-.17	-.13
Emotional Stability	-.45*	-.17	1	.04
Extraversion	-.02	-.13	.04	1
Openness	-.15	.22	.08	-.06
BIS/BAS				
Behavioural Inhibition	.38*	.07	-.73***	.12
Drive	.38	-.20	-.41*	-.18
Fun	.17	.10	-.22	-.42*
Reward	.17	.02	-.44*	.30
Importance trust	.02	-.01	.06	.39

	Openness	Behavioural Inhibition	Drive	Fun
PANAS				
Positive Affect pre pain	.46*	.22	-.27	-.26
Negative Affect pre pain	-.08	-.08	-.34	-.42
Positive Affect post pain	.42*	.43	-.18	-.20
Negative affect post pain	-.06	.01	-.53***	-.25
STAI				
Trait Anxiety	-.06	-.64***	-.13	-.02
State Anxiety	-.12	-.43*	-.27	-.14
FPQ				
Fear of Pain	-.13	-.42*	-.004	.02
PCS				
Pain Catastrophizing	.11	-.39*	-.26	-.30
BFI				
Agreeableness	-.15	.38	.38*	.17
Conscientiousness	.22	.07	-.02	.10
Emotional Stability	.08	-.73***	-.41*	-.22
Extraversion	-.06	.12	-.18	-.42
Openness	1	.02	-.10	-.21
BIS/BAS				
Behavioural Inhibition	.02	1	.17	-.04*
Drive	-.10	.17	1	.68
Fun	-.21	-.04	.68***	1
Reward	-.23	.13	.67***	.70***
Importance trust	-.05	-.12	-.49**	-.45

	Reward	Importance trust
PANAS		
Positive Affect pre pain	-.13	.11
Negative Affect pre pain	-.37	.23
Positive Affect post pain	-.08	-.14
Negative affect post pain	-.37	.16
STAI		
Trait Anxiety	-.19	-.16
State Anxiety	-.32	-.03
FPQ		
Fear of Pain	-.01	.14
PCS		
Pain Catastrophizing	-.22	.20
BFI		
Agreeableness	.17	.02
Conscientiousness	.02	-.009
Emotional Stability	-.44	.06
Extraversion	-.30	.39*
Openness	-.23	-.05
BIS/BAS		
Behavioural Inhibition	.13	-.12
Drive	.67	-.49**
Fun	.70***	-.45
Reward	1	-.50**
Importance trust	-.50**	1

Note. *$p<.05$, ** $p<.01$, *** $p<.001$