Revealing the latent structure of multidimensional facial features that drive social trait perceptions

Laura B. Hensel

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

School of Psychology and Neuroscience
College of Science and Engineering
University of Glasgow

March 2022

Abstract

Humans readily attribute social traits to others based on their facial appearance, influencing behaviors, social interactions, and decision making. According to influential models of social perception, these judgments are based on two central dimensions – trustworthiness/warmth and dominance/competence. Because of the wide-reaching impact of such social attributions, a long-standing focus has been to identify the facial features that elicit these social judgments. However, the face is complex, comprising 3D shape, 2D complexion, and dynamic facial expressions, making such investigations empirically challenging. As a result, central questions regarding fundamental facial features of social traits, and how these relate to other social judgments such as social class and emotion, remain unanswered. In this thesis, I use data-driven psychophysical methods to mathematically model the 3D shape, 2D complexion and dynamic facial expressions that elicit judgments of four key social trait dimensions – dominance, competence, trustworthiness, and warmth. I then identify the latent face feature space underlying these judgments using a data-reduction technique. Results reveal two latent 3D shape and three latent 2D complexion feature spaces on the basis of which social traits cluster into four distinct subgroups. Moreover, I show that these social trait feature spaces correlate positively with facial expression features (shape) and age-cues (complexion). I then examine whether these features also drive perceptions of another important social judgment – social class. To do so, I model the 3D shape and 2D complexion of social class, using the same approach as for social traits. I compare these models using a data reduction technique and a supervised machine learning approach. Results show that in line with conceptual overlaps arising from social class stereotypes (e.g., poor = incompetent), social class and social trait dimensions share facial features. However, no single trait's features fully account for social class features. Finally, I compare the social trait facial expression models to emotional expressions and social trait face shapes to reveal the features shared between each. Results showed that longstanding associations between perception of emotion and social traits (e.g., happy = trustworthy), only partially account for social trait perception. The current work informs central theories of social perception, highlights drivers of socially relevant stereotype associations, and can aid the development of psychologically grounded digital agents.

Contents

List of Tables

List of Figures

Publications

Parts of this thesis have been published as peer-reviewed conference abstracts and proceedings:

Chapter 2

Hensel, L. B., Zhan, J., Bjornsdottir, R. T., Garrod, O. G., Schyns, P. G., & Jack, R. E. (2020). Psychologically Valid Social Face Features for Virtual Agents. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents (pp. 1-3)*, doi: https://doi.org/10.1145/3383652.3423899.

Hensel, L. B., Zhan, J., Bjornsdottir, R. T., Garrod, O. G., Schyns, P. G., & Jack, R. E. (2020). Social trait perception is structured by a latent composition of 3D face features. *Journal of Vision, 20*(11), 1365-1365, doi: https://doi.org/10.1167/jov.20.11.1365.

Chapter 4

Hensel, L. B., Garrod, O. G., Schyns, P. G., & Jack, R. E. (2021). Facial expressions of social traits comprise core affective signals; Top Ranked Abstracts from the 2021 Annual Meeting of the Society for Affective Science. *Affective Science*. https://doi.org/10.1007/s42761-021-00070-w.

Hensel, L. B., Garrod, O. G., Schyns, P. G., & Jack, R. E. (2021). Social Trait Facial Expressions Comprise Latent Affective Facial Signals. *Journal of Vision, 21*(9), 1988-1988, doi: https://doi.org/10.1167/jov.21.9.1988.

Acknowledgments

The work I present in this thesis and other research activities I had the opportunity to complete in the last three and a half years would not have been possible without the support of many insightful, kind, and knowledgeable people. I would first like to thank Prof. Rachael E. Jack for her support and invaluable advice at every stage of my PhD and the research projects within it. I would also like to thank Prof. Philippe G. Schyns, Dr. Jiayu Zhan, and Dr. Robin A. A. Ince for their indispensable input and for laying the groundwork for many of these projects. I equally owe much thanks to Dr. Oliver G.B. Garrod for his work on the face generation platform that made it possible for me to include 963 faces in this thesis without which producing figures would have been only half the fun. My immeasurable thanks also to Dr. R. Thora Bjornsdottir for not only being a magnificent collaborator but also a wonderful friend. Thank you to my annual progression reviewers Prof. Lisa DeBruine and Dr. Robin Ince. Your advice never led me astray and your fresh views on my work were always immensely helpful. I would also like to thank other members of the lab who have made this time much easier and more enjoyable through their scientific and personal advice. In particular Valentina Gosetti, Meng Liu, and Chaona Chen. Lastly, I would like to thank the Scottish Graduate School of Social Science for the studentship that allowed me to conduct this thesis.

An enormous thank you as well to my parents and sister. You provided me with the foundations that got me here and your belief in me has always given me strength. Thank you, Franzi, for being the non-Psychology academic that I needed to balance my life, my all-important gym buddy, and first and foremost, an amazing friend. Finally, I owe infinite thanks to Jack for his love, unwavering emotional support, patience, and belief in me. You were the rock that everyone needs to get through a PhD and I will always owe you for that.

Declaration

With the exception of chapter 1, which contains introductory material, all work in this thesis was carried out by the author unless otherwise explicitly stated. A number of chapters have been conducted in conjunction with co-authors, in particular chapter 3. In all cases, the author has held a leading role in the project and made the primary contribution to the work as presented here. This work has not been submitted in any previous application for a higher degree or professional qualification.

Signature:

List of Abbreviations

d Euclidean Distance

SD Standard Deviation

AU Facial Action Unit

fWHR Facial Width-to-Height Ratio

GFG Generative Face Grammar

GLM General Linear Model

GLME Generalized Linear Mixed Effects Model

GMF Generative Model of 3D Face Identity

GUI Graphical User Interface

LOOCV Leave-One-Out Cross Validation

MI Mutual Information

PC Principal Component

PCA Principal Component Analysis

SVM Support Vector Machine

Chapter 1

General introduction

A central part of the human experience are the impressions we form of others and others form of us. We build and up-date these impressions frequently (e.g., Jaeger et al., 2019; Miller et al., 2004), extremely quickly (Todorov et al., 2009; Willis & Todorov, 2006), and they influence our social interactions and behaviors (for a review see Todorov et al., 2015). The outcomes of these first impressions range from the seemingly trivial – such as engaging with someone at a party – to the highly important – such as hiring decisions (Varghese et al., 2018). A central aim within psychology has therefore been to identify the precise antecedents of such social judgments. To do so, one has to first understand a) which social traits are fundamental to how humans view each other and b) the precise characteristics that elicit perceptions of these social traits. A large body of research has devoted itself to answering the first of these questions: Identifying the social traits central to human perception. Similarly, identifying the precise characteristics that drive judgments of these traits has also garnered much attention. However, it is the infinitely more complicated question to answer because humans base their impressions on a multitude of appearance and behavioral cues (Campanella & Belin, 2007; Hehman et al., 2015; Quadflieg & Westmoreland, 2019; Todorov et al., 2015), not to mention their own disposition and knowledge (Xie et al., 2019). The current work focuses therefore on a single part of the human body – the face – which is not only central to impression formation (Jaeger et al., 2019; Zebrowitz, 2018) but also to human social interaction (Jack & Schyns, 2017). The face comprises two intrinsic cues – face shape and face complexion (i.e., skin color, blemishes, wrinkles) – and is additionally capable of producing facial expressions for social communication. This thesis examines how each of these three components of the face elicit perceptions for four key social trait dimensions – dominance, competence, trustworthiness, and warmth – to a) identify the facial features (shape, complexion, and facial expressions) that are fundamental to social trait perception and b) explore how these face features and facial expressions relate to face features of other socially relevant judgments.

1.1 Prevalence and impact of social trait perceptions

First impressions permeate our daily lives with children as young as three years old inferring social traits from faces (Cogsdill et al., 2014). Such early judgments are socially reinforced by parents (Eggleston et al., 2021) and continue throughout our lifetime (Connolly et al., 2021). Consequently, humans are extremely proficient at attributing social traits to others and make reliable judgments after observing a face for just 100 milliseconds (Albert et al., 2021; Willis & Todorov, 2006). Such attributions form automatically (Ritchie et al., 2017), subconsciously (Janssens et al., 2020), and from minimal cues (Bacev-Giles & Haji, 2017). Moreover, humans frequently rely on facial social trait cues for decision making even when more diagnostic information, such as others' motivations, is available (Jaeger et al., 2019). This is likely because facial information is readily available and intuitively processed (Albert et al., 2021). Indeed, even caricatures of faces and facial expressions, such as emoticons, lead to impression formation (Glikson et al., 2018) and humans also assign such social traits to non-human entities (DeSteno et al., 2012). Similarly, people readily form first impressions based on others' social media profiles (Bacev-Giles & Haji, 2017). Importantly, such first impressions can be long-lasting. For example, reliable information about someone's trustworthiness fails to change initial judgments made based on that person's facial appearance unless the new information is perceived to be extreme and reliable (Shen & Ferguson, 2021). Especially initial negative impressions are long-lasting and resistant to change (Baumeister et al., 2001; Richey et al., 1967). Crucially, these quickly formed and enduring first impressions affect many aspects of human life.

Such influence of first impressions ranges from from romantic choices to hiring decisions (for a review see Antonakis & Eubanks, 2017). For example, social trait perception influences partner choices with a general preference for warm and trustworthy partners (Valentine et al., 2020). Especially in the selection of long-term partners, perceived trustworthiness plays a major role (Carrito et al., 2020), while attractiveness is more important in short initial meetings (Fletcher et al., 2014; Valentine et al., 2013). Similarly, political candidates' facial appearance, including their perceived trustworthiness, babyfacedness (i.e., facial features that resemble those of a baby), competence, and attractiveness, affects their election success (Ahler et al., 2016; Franklin & Zebrowitz, 2016; Joo et al., 2015; Sussman et al., 2013; Todorov et al., 2005). In particular competence judgments play an important role in such decisions (Olivola & Todorov, 2010). Similarly, social trait judgments influence hiring choices (Cuddy et al., 2011; Rudman & Glick, 1999). For example, perceived facial competence predicts both hiring and compensation of CEOs (Graham et al., 2012) and perceived facial trustworthiness correlates with managers' positions in corporate hierarchy (Linke et al., 2016). Social trait judgments furthermore affect social interactions (Chua & Freeman, 2022; Qi et al., 2021; Shang & Li, 2020). For example, in economic games participants are more likely to punish untrustworthy looking faces (Shang & Li, 2020), allocate more payment to trustworthy looking faces (Chua & Freeman, 2022), and are more willing to take risks after viewing a trustworthy face (Qi et al., 2021). Moreover, facial competence and warmth impacts whether others perceive one's social exclusion as damaging (Rudert et al., 2017). In sum, social trait inferences have wide-reaching influences on human life and have thus long been a central focus in social psychology.

1.1.1 Judgment accuracy

Despite such wide influence of social attributions, social trait judgments are not always *accurate* (e.g., Todorov et al., 2015). That is, the social traits humans attribute to others based on their appearance seldom predict actual behavior or social outcomes (Todorov et al., 2015). Although there is a deluge of work showing the apparent predictive power of facial social trait ratings for real world behavior (e.g., Mueller & Mazur, 1996; Porter et al., 2008), these are frequently marred by methodological issues such as poor control variables (Todorov et al., 2015). For example, facial trustworthiness was shown to predict criminal behavior (Porter et al., 2008). However, when controlling for possible confounding factors such as age, ethnicity, and gender, the link between trustworthiness judgments and a number of behavioral markers, including criminality and willingness to cheat on a test, disappeared (Rule et al., 2013). Similarly, while perceived facial competence does predict CEO hiring and compensation (Graham et al., 2012), there is no link between CEOs' facial competence and actual company performance (Stoker et al., 2016). Instead, already profitable companies hire CEOs that look more competent (Graham et al., 2012), indicating that one's facial appearance is likely to predict *others'* behavior towards oneself rather than one's own.

Indeed, many of the links between facial appearance and real-world behavior are likely due to self-fulfilling prophecy effects (Zebrowitz, 2017, 2018). For example, individuals with higher perceived facial dominance tend to hold higher military rank (Mueller & Mazur, 1996). However, when considering why this might be, one has to take into account stereotypes about dominant-looking individuals (Haselhuhn & Wong, 2012; Stirrat & Perrett, 2010) that likely impact others' behavior towards them (Haselhuhn et al., 2013), such as that they are risk-takers (Hareli et al., 2021). Consequently, as with the example of the competent looking CEOs, dominant-looking individuals likely get selected for particular leadership roles (Olivola et al., 2014). Additionally, any links between appearance and behavior are compounded by individuals' awareness of their own appearance (Slepian & Ames, 2016). That is, people are aware of how they are perceived by others and act accordingly. For example, social media profiles elicit accurate social attributions (e.g., Back et al., 2010) as they allow individuals to strategically present information about themselves (Todorov et al., 2015). However, generally human raters are more accurate when basing their judgments on behavioral rather than appearance cues (Olivola & Todorov, 2010).

That is not to say, however, that social trait judgments are *never* accurate. Indeed, this would be counter-intuitive as accurate social judgments would yield an evolutionary benefit (Zebrowitz, 2018). In fact, a recent meta-analysis found that trustworthiness impressions are indeed somewhat accurate with small to moderate effects (Foo et al., 2021). Some behavioral markers stood out in particular as being accurately predicted by facial trustworthiness: aggressiveness and unfaithfulness (Foo et al., 2021). This suggests that trustworthiness judgments do meaningfully reflect certain behaviors based on physical correlates. Correspondingly, both aggression and sexual faithfulness are linked to testosterone levels (Arnocky et al., 2018; Carre et al., 2013; Eisenegger et al., 2011) which in turn are associated with facial appearance (Lefevre et al., 2013; Penton-Voak & Chen, 2004), explaining why facial appearance more accurately predicts these behaviors. In contrast, other contexts reduce judgment accuracy. For example, domain-specific experts have particularly poor accuracy when forming first impressions of others in the context of their specialist domain (Re & Rule, 2016). Additionally, some social attributes, such as competence, are judged with lower accuracy than others and their judgment accuracy is further affected by other factors, including age (Zebrowitz et al., 2014). Accuracy of social judgments is therefore dependent not only on the social situation and behavioral correlate in question but also the attribute itself.

Despite these mixed findings for social trait judgment accuracy, raters over-estimate their ability to accurately judge others' social traits (Ames et al., 2010; Hassin & Trope, 2000; Olivola et al., 2014). Such meta-accuracy, i.e., how accurately someone judges their own performance on a social attribution task, is less well studied than accuracy. However, findings suggest that people who believe they are good at judging others are no better in terms of social trait judgment accuracy than people who are less confident in their social trait judgments (Olivola et al., 2014). Especially in light of the wide-reaching impact of social judgments (see section 1.1), such over-estimation of ones' performance is worrying. Perhaps such poor meta-accuracy is in part due to the judgment consistency between raters. That is, while social trait judgments may not be accurate, raters frequently show high agreement in their judgments (e.g., South Palomares & Young, 2018; Sutherland et al., 2018; Vernon et al., 2014; Walker & Vetter, 2009; Willis & Todorov, 2006). Such inter-rater agreement may be falsely interpreted by individuals as accuracy. Together, this work suggests that while social trait judgments are highly reliable, they are only accurate in some contexts.

1.2 Theories and models of social trait perception

To make sense of how these social trait judgments are formed, a number of theories of social perception emerged. Such theories are commonly concerned with one or several of three aspects of social perception: impression accuracy, consensus, and functionality (Quadflieg & Westmoreland, 2019). One of the earliest theories that was particularly focused on perceptual accuracy, the Lens Model, proposed that stimuli comprise multiple cues that are imperfectly correlated with one another (Brunswik, 1956). In this framework, the perceiver acts as an intuitive statistician to combine these cues to arrive at a judgment with some accuracy (Brunswik, 1956). Later, the model was refined for social inferences (Scherer, 1978, 2003) and successfully applied to impression formation (Gifford, 1994). However, the Lens Model is inherently correlational and assumes that perceivers base their judgments of a given social trait on accurate weightings of different cues (Zebrowitz, 2018) and has thus not remained popular in the social perception literature.

Later theories of social perception expanded the view of observers as intuitive statisticians and moved away from the focus on impression accuracy. For example, the Weighted Average Model of social perception posits that observers base their social trait judgments on several behavioral cues each with a given weight (Carroll & Anderson, 1982; D. A. Kenny, 1991). It is therefore similar to the Lens Model in that it views the perceiver as an intuitive statistician and comes with similar limitations. However, the Weighted Average Model acknowledges that the weights perceivers assign to different factors are inherently subjective and unique to each person (D. A. Kenny, 1991). Similarly, the Stage Model of Dispositional Inferences sees impression formation as a goal-directed and iterative behavior (Trope & Higgins, 1993). Within this framework, perceivers form impressions either algorithmically – similar to the Lens Model – or heuristically, informed and biased by the perceiver's own dispositions and experiences – similar to the Weighted Average Model (Trope & Higgins, 1993). However, in contrast to previous theories, the Stage Model posits that perceivers continually update these algorithmically or heuristically derived impressions based on their observed validity (Trope & Higgins, 1993). The Stage Model therefore views social perceptions as a dynamic process, making this model akin to more recent theories of social perception.

1.2.1 Recent theories of social trait perception

Recent theories of social perception acknowledge the complex nature of social perception. For example, Dynamic Interactive Theory posits that bottom-up sensory cues and top-down cognitive influences interact to determine first impressions (Freeman & Ambady, 2011; Freeman et al., 2020). Specifically, the theory proposes that social trait perceptions emerge from a complex integration of vast amounts of bottom-up and top-down information.

Particularly, auditory and visual sensory cues interact with category level cues such as sex or ethnicity that in turn interact with stereotypes and higher-level cognitive states such as motivation or task demands (Freeman & Ambady, 2011). For instance, the visual cue of masculine facial features might lead to an inference of male gender which in turn is associated with stereotypes of aggression (Boshyan et al., 2014; Geniole et al., 2014). As within the Stage Model, Dynamic Interactive Theory posits that these inferences are continually updated through successive cycles of interaction (Freeman & Ambady, 2011). The theory therefore shares similarities with previous frameworks and further refines our understanding of human perceivers as complex systems.

The theories discussed above are concerned with how different cues integrate to form first impressions and therefore focus primarily on *how* humans arrive at social judgments. Two further theories place a greater emphasis on *why* social perceptions arise. Specifically, the Social Functional approach proposes that social perceptions hold some evolutionary benefit – they serve a social function (Adams et al., 2017). A human observer makes behavioral forecasts based on sensory cues and these forecasts should ideally have some ecological validity (Adams et al., 2017). In particular behaviors such as approach and avoidance and identifying threat play an important evolutionary role (Adams et al., 2017). Similarly, the Ecological Theory of social perception emphasizes the evolutionary benefits of accurate social perceptions, thereby focusing on how sensory cues lead to social inferences (McArthur & Baron, 1983; Zebrowitz, 2017). These theories therefore focus primarily on the perceptual basis of social inferences – which also forms the focus of this thesis. Particularly the Ecological Theory of social perception has been central in the study of how humans perceive social traits from faces (Knutson, 1996; McArthur & Baron, 1983; Said et al., 2009; Zebrowitz, 2017).

1.2.2 Ecological theory of social trait perception

Ecological Theory was first developed as a general account of visual perception (Gibson, 1979) and later adapted to social perception specifically (McArthur & Baron, 1983). The theory posits that social perceptions serve an adaptive function (McArthur & Baron, 1983). That is, observers aim to focus on *useful* information contained in the complex multimodal stimuli they encounter (McArthur & Baron, 1983). Different observers may be attuned to different sources of information (McArthur & Baron, 1983) and thus make different perceptual judgments. The theory thereby captures a potential reason for why humans make social inferences while also explaining judgment variance between perceivers. For example, humans are attuned to recognizing physical attributes of infancy because appropriate behavior around infants yields an adaptive advantage (McArthur & Baron, 1983; Zebrowitz, 2017). At face value, Ecological Theory therefore contradicts the low social trait judgment accuracy discussed in subsection 1.1.1. If the attribution of social traits from facial features serves adaptive value, then it is puzzling why humans perform relatively poorly at this task. However, Ecological Theory provides a solution for this contradiction: it proposes that high adaptive attunement comes at the cost of overgeneralization (McArthur & Baron, 1983). In the case of infant features, for example, observers ascribe the social attributes associated with infancy not only to babies but also to adults that exhibit infant-like physical features such as round eyes and a round face (i.e., 'babyface'; Berry & McArthur, 1985; Zebrowitz et al., 2003). In support of this hypothesis, babyfaced individuals are perceived as trustworthy, warm, submissive, and incompetent (Jaeger et al., 2020; Oosterhof & Todorov, 2008; Zebrowitz et al., 2003; Zebrowitz et al., 2011).

Another important overgeneralization effect is emotion overvegeneralization. Emotion overgeneralization refers to two related effects: (1) emotions are reliably inferred from purportedly neutral faces, an effect that was documented as early as the 1980s (Malatesta et al., 1987), and the reverse effect (2) stable social traits are inferred from transient emotional facial expressions (Knutson, 1996). The former of these effects – emotion overgeneralization from neutral faces – is likely driven by neutral face features that resemble emotional expressions (Zebrowitz, 2017). For example, neutral faces that objectively resemble a happy facial expression are judged as more likable, less dangerous, and more trustworthy than those resembling an angry facial expression (Adams et al., 2012; Oosterhof & Todorov, 2009; Zebrowitz et al., 2010). Indeed, emotional resemblance is a central driver for impression formation and emotional facial content predicts trustworthiness and dominance judgments better than other facial attributes (Jaeger & Jones, 2022).

Emotion overgeneralization is further influenced by face gender (Albohn & Adams, 2021; Zebrowitz et al., 2010). For example, female faces generally appear more infant-like and comprise babyface features such as large eyes (Friedman & Zebrowitz, 1992; McArthur & Apatow, 1984). In turn, a widening of the eyes forms part of the surprise facial expression (Ekman & Friesen, 1978; Jack et al., 2012). Female faces therefore physically resemble surprise facial expressions and consequently elicit higher surprise judgments than male faces (Zebrowitz et al., 2010). Similarly, male faces physically resemble angry facial expressions and both male and angry looking faces are rated as dominant (Albohn & Adams, 2021). Overall, the overgenerlization literature placed a strong focus on face gender, often in conjunction with babyfacedness (Albohn & Adams, 2021; Hess et al., 2009; Hess et al., 2000; Zebrowitz et al., 2010). However, others have noted that emotion overgeneralization effects are indeed independent of babyfacedness (Montepare & Dobish, 2003). Importantly, masculine looking female faces also elicit higher dominance judgments than feminine female faces (Albert et al., 2021; Sutherland et al., 2015), suggesting it is the specific facial feature composition, rather than gender as a construct, which drives emotion overgeneralization

(Said et al., 2009).

The emotion overgeneralization effect extends such that observers infer social traits from emotional facial expressions. Much as neutral faces that appear angry are judged as dominant, and those that appear happy as trustworthy, so too are angry expressions judged as dominant, and happy expressions as trustworthy (Hareli et al., 2009; Knutson, 1996). Interestingly, variance in social judgments based on emotional expressions is lower compared to those based on neutral faces (Hehman et al., 2017), suggesting that the more salient emotional expression leads to greater agreement among raters. Considering the evidence for shared physical features between social traits and emotions (Albohn & Adams, 2021; Jaeger & Jones, 2022; Said et al., 2009), it may be the exaggeration of those physical features that enables this greater judgment reliability. However, it is unclear whether facial expressions are mere exaggerations of neutral face features and both signals are degenerate – signaling the same information (i.e., social traits) – or whether facial features that convey social traits comprise additional featural components that are not linked to emotion perception.

1.2.3 Dimensional models of social trait perception

The development of these theories was accompanied by a drive to identify the main dimensions of social perception (Allport & Odbert, 1936; Asch, 1946; Ashmore & Tumia, 1980; Fiske et al., 2002; Oosterhof & Todorov, 2008). First, the central importance of social attributions to human life (e.g., Albert et al., 2021; Antonakis & Eubanks, 2017; Jaeger et al., 2019, see also section 1.1) and later the predictions of evolutionary and functional theories that social perceptions are shaped by key social dimensions (Adams et al., 2017; McArthur & Baron, 1983), lead to this push to find the main dimensions along which such attributions are made. Systematic collations of trait descriptors date back as far as the 1930s (Allport & Odbert, 1936). Early work on impression formation was not, however, focused on dimensional models. For example, Solomon Asch, a pioneer of impression research, held a Gestalt, not a dimensional, view of trait impressions (Asch, 1946). That is, he viewed impressions as formed on the basis of several traits which interact and add up to a single unified and coherent view of a person (Asch, 1946). Importantly, in this view, the contribution of any given social trait to impression formation is affected by other traits associated with the same person. In his seminal work, Asch (1946) demonstrated that some social traits, such as warmth, are central to first impressions. However, he also showed the traits' interdependence and demonstrated, for example, that aggressiveness can be seen as friendly and open when presented together with adjectives such as active and helpful (Asch, 1946). However, social psychology subsequently embraced the dimensional over the Gestalt approach and frequently clustered trait descriptors to derive primary dimensions of social perception. For example, Rosenberg et al. (1968) used this technique to identify three central dimensions for impression formation: good-bad (or social desirability), hard-soft (or intellectual desirability), and active-passive. These bear resemblance to the main dimensions Ashmore and Tumia (1980) identified in their investigation of sex stereotypes: social desirability, intellectual desirability, and soft-hard which correlated highly with female-male. Together, this body of work laid the foundation for modern dimensional models of impression formation.

1.2.4 The stereotype content model

A highly influential model of first impressions, the stereotype content model, comes from the field of person perception (Fiske et al., 2002; Fiske et al., 2007). As the name suggests, this model is primarily concerned with how humans form stereotypes of others. Fiske et al. (2002) propose that, when meeting others, people primarily focus on two things: Firstly, whether this person means them harm and secondly, whether the person has the ability to act on their possible intent for harm. This focus on social perception as a functional process associated with approach and threat perception make this model very similar to the Social Functional and Ecological theories (Adams et al., 2017; McArthur & Baron, 1983). Within the stereotype content model, these concepts are captured by the warmth (intent for harm) and competence (ability) dimensions (Fiske et al., 2002; Fiske et al., 2007). According to Fiske, these two dimensions together make up perceptions of stereotyped groups. For example, Fiske et al. (2002) demonstrated that U.S. Americans view Asians as high in competence but relatively low in warmth, which is likely due to their competence being perceived as threatening (M. H. Lin et al., 2005). More recent work demonstrated that people mentally represent facial features of stereotyped groups, such as nurses and managers, as reflecting their respective warmth and competence (Imhoff et al., 2013). Warmth and competence are also important dimensions cross-culturally (Sutherland et al., 2018). For instance, for both British and Chinese participants, warmth emerged as the primary dimension followed by competence when reporting spontaneous first impressions (Abele et al., 2016; Sutherland et al., 2018).

This and previous work (Fiske et al., 2007) point towards a primacy-of-warmth effect wherein warmth is the first and more important dimension to be evaluated. Specifically, participants rely on warmth-related concepts, such as trustworthiness and sincerity, more so than competence-related traits, such as intelligence, when forming first impressions of strangers (Brambilla et al., 2011; Wayne Leach et al., 2007) and when updating these impressions after receiving new information (Luttrell et al., 2022). However, there is some evidence that the primacy-of-warmth effect is context dependent with aspects of competence, such as intelligence, also playing important roles (Abele & Wojciszke, 2007; Nauts et al., 2014). Additionally, there is some debate over the exact relationship between warmth and competence. While the two traits are traditionally seen as orthogonal (Fiske et al., 2002), others have suggested a curvilinear relationship whereby competence perceptions increase with warmth perception up to the point of highest warmth after which perceptions of competence decline (Imhoff & Koch, 2017). Similarly, the experimental task, i.e., whether participants compare two groups or judge each separately, also affects the relationship between warmth and competence (Judd et al., 2019; Judd et al., 2005). This relationship is further complicated by the fact that warmth itself comprises two empirically distinct dimensions: morality and sociability (Abele et al., 2016; Brambilla & Leach, 2014; Goodwin, 2015). Morality encompasses traits such as trustworthiness and sincerity while sociability refers to warmth, friendliness, and likability (Brambilla & Leach, 2014; Brambilla et al., 2011; Luttrell et al., 2022).

Despite these theoretical debates, it is clear that the stereotype content model plays a central role in person perception (Fiske, 2018). For example, important interpersonal attitudes, such as (dis)like and respect, are tightly linked to warmth and competence judgments (Wojciszke et al., 2009). Specifically, warmth perception influences like and dislike while perceived competence predicts respect (Wojciszke et al., 2009). Importantly, although the stereotype content model originated from the person perception, rather than face perception literature, observers readily perceive warmth and competence from face information (Imhoff et al., 2013; Kervyn et al., 2015; Sutherland et al., 2016; Wen et al., 2020). For example, feminine-looking faces, regardless of gender, are seen as warmer than masculine-looking faces which in turn are perceived as more competent (Wen et al., 2020). Additionally, warmth and competence relate to threat and status perception (Fiske et al., 2002; Kervyn et al., 2015). Specifically, faces perceived as warm appear less threatening than colder looking faces (Kervyn et al., 2015). Similarly, competence perceptions correlate positively with perceived status (Fiske et al., 2002). This, together with the cross-cultural evidence for the stereotype content model (Abele et al., 2016; Sutherland et al., 2018), make this one of the central models for social trait perception.

1.2.5 The trustworthiness-dominance model

In contrast to the primarily top-down, theoretically derived and later experimentally validated stereotype content model, the trustworthy-dominance model arose from bottom-up clustering of face judgments (Oosterhof & Todorov, 2008). Specifically, participants provided unconstrained descriptions of faces which yielded 14 most frequent trait descriptors. A new set of participants then rated the faces again on these 14 social traits with the addition of dominance. Finally, Oosterhof and Todorov (2008) used Principal Component Analysis (PCA) to derive the main dimensions of face evaluation from these judgments. This resulted in two Principal Components (PCs) that, together, accounted for 81.6% of the variance in face judgments and that correlated with trustworthiness and dominance attributions, respectively (Oosterhof & Todorov, 2008). The authors therefore derived a two-factor model of social trait perception from faces with two orthogonal dimensions: trustworthiness and dominance. Notably, dominance did not emerge from the data as one of the most frequently used descriptors but was included by the authors due to its theoretical relevance (Oosterhof & Todorov, 2008; Wiggins, 1979). Later work, based on a similar design, did not identify a dominance dimension but did identify both a trustworthiness/warmth and a competence dimension (Sutherland et al., 2018).

Notably, the trustworthiness-dominance model resembles the stereotype content model closely. The trustworthiness and warmth dimensions are conceptually similar, referring to intent (e.g., Fiske et al., 2007), although others propose that they are separate dimensions of morality (trustworthiness) and sociability (warmth) (Brambilla & Leach, 2014; Goodwin, 2015). Similarly, the dominance and competence dimensions both relate to ability (e.g., Fiske et al., 2002; Walker & Vetter, 2016). However, while trustworthiness and warmth judgments correlate highly, dominance and competence share only low perceptual similarity (Oliveira et al., 2019; Sutherland et al., 2016; Sutherland et al., 2020). In fact, Oliveira et al. (2019) found only a small correlation between pixel luminance values of competent and dominant faces and almost no correlation between incompetent and submissive faces. Additionally, competence perceptions are more variable and depend, more so than dominance perceptions, on perceiver characteristics (Hehman et al., 2017). It is, however, poorly understood which facial features drive these perceptual differences. What is more, similar to the primacy-of-warmth effect and Asch's (1946) idea that social traits carry different weights, trustworthiness and dominance, too, do not combine linearly during impression formation (Oliveira et al., 2019). Specifically, Oliveira et al. (2019) used reverse correlation (see subsection 1.4.1) to derive classification images of faces that represented participants' mental representations of a specific *combination* of trustworthiness and dominance. Analysis of the resulting classification images of each trustworthiness-dominance combination revealed that trustworthiness and dominance were not weighted equally and, frequently, trustworthiness took precedence over dominance (Oliveira et al., 2019), supporting the primacy of warmth hypothesis.

Nevertheless, both trustworthiness and dominance play important roles in social trait perception from faces. Trustworthiness and dominance judgments occur extremely quickly (Albert et al., 2021; Todorov et al., 2009; Willis & Todorov, 2006) and subconsciously (Janssens et al., 2020; Ritchie et al., 2017). For example, participants rated masculinized faces as dominant after just 100 milliseconds of exposure (Albert et al., 2021). Similarly, trustworthiness perceptions can bypass the early visual cortex which is involved in conscious perception (Janssens et al., 2020), suggesting automatic and subconscious processing of these social traits. Indeed, perception of trustworthiness from facial information endures even for richer stimuli, i.e., stimuli in which more than the facial information is available (Rezlescu et al., 2012). Even in neutral contexts, trustworthiness of a face is perceived spontaneously by observers (Klapper et al., 2016). Trustworthiness and dominance are therefore readily perceived from human faces and form fundamental dimensions of social perception that are similar to the warmth and competence dimensions.

1.2.6 Central social dimensions - conflicting findings

The stereotype content model and the trustworthiness-dominance model both posit that social trait perception is based on two underlying dimensions which roughly correspond to intent and ability (Fiske et al., 2002; Oosterhof & Todorov, 2008). However, the precise number of central social dimensions remains unclear. For example, a large registered report which included data from eleven world regions failed to reliably replicate the trustworthiness-dominance model (B. C. Jones et al., 2021). Even in Western countries, within which the model was first developed, B. C. Jones et al. (2021) identified more than two factors of social attributions. Similarly, work based on ambient images (Jenkins et al., 2011) identified three rather than two fundamental dimensions: an approachability and dominance dimension, highly similar to those in the work of Oosterhof and Todorov (2008), and an additional youthfulness-attractiveness dimension (Sutherland et al., 2013; Vernon et al., 2014). Importantly, ambient images vary in a wide range of aspects such as age, health, facial hair, and head angle (Jenkins et al., 2011). In contrast, the highly constrained nature of the original images (Oosterhof & Todorov, 2008) potentially meant that aspects such as age did not vary sufficiently to emerge as unique dimensions. It should also be noted that youthfulness-attractiveness correlated with the trustworthiness dimension (Sutherland et al., 2018), indicating that the two dimensions are not independent.

Indeed, perceptions of age correlate highly with several social trait dimensions. For example, older aged adults are seen as less competent than younger adults (Cuddy et al., 2005), a trend that is not limited to Western cultures and which is exacerbated by aging populations (Berry & McArthur, 1985; North & Fiske, 2015). Moreover, observers more accurately judge competence from younger compared to older adult faces (Zebrowitz et al., 2014), suggesting a perceptual bias that compounds these age-related stereotypes. Similarly, facial cues to perceived and actual health (Henderson et al., 2016) influence social trait and maturity perception (Jaeger et al., 2018; Tsankova & Kappas, 2016). For example, blemishes not only decrease perceptions of health and age but also of trustworthiness and competence (Jaeger et al., 2018).

Subsequent work, which focused on clustering faces rather than using dimension reduction, instead identified a single dimension – approachability (i.e., approach-avoidance; A. L. Jones & Kramer, 2021). This central approachability dimension aligns with the primacy-of-warmth hypothesis (Fiske et al., 2007) and the fact that the trustworthiness-related dimension frequently explains the majority of variance in judgments (Oosterhof & Todorov, 2008; Sutherland et al., 2018; Sutherland et al., 2013; Vernon et al., 2014). Together, this suggests that approach-avoidance is the primary dimension of social perception and corresponds to warmth and trustworthiness (Slepian et al., 2017; Todorov, 2008). However, the studies discussed so far were based on a relatively small amount of social trait descriptors – commonly between twelve and sixteen (Oosterhof & Todorov, 2008; Sutherland et al., 2018; Vernon et al., 2014). In contrast, analysis based on 100 descriptors C. Lin et al. (2021) identified four main dimensions: warmth, competence, femininity, and youth. These dimensions align with those identified by Sutherland et al. (2018), Sutherland et al. (2013) but include the additional dimension of femininity. Interestingly, femininity and masculinity have recently been proposed to underlie social perceptions and specifically all two-factor models, including the stereotype content and trustworthiness-dominance models (A. E. Martin & Slepian, 2021). This theory is tentatively supported by evidence that femininity and masculinity form two separate dimensions rather than representing the ends of a single dimension (Hester et al., 2021). However, aside from this recent development, most of the work investigating central dimensions of social perceptions has identified social traits which correspond loosely to the dimensions of intent and ability although there is some variability and potentially additional dimensions (e.g., Fiske et al., 2002; Fiske et al., 2007; C. Lin et al., 2021; Oosterhof & Todorov, 2008; Sutherland et al., 2018; Vernon et al., 2014). As a result, the current work focuses on these four central social trait dimensions: dominance, competence, trustworthiness, and warmth.

1.2.7 Central social dimensions - social class

Perceptions of these four social traits correlate with another important social judgment: Social class. Social class forms a central hierarchy across many species (e.g., Chiao, 2010; Sapolsky, 2004). In human society, socioeconomic status in particular is central (e.g., Kraus et al., 2013) and, as with social traits, perceptions of social status have important social ramifications (e.g., Adler et al., 2000; Morrison, 2019; Nishi et al., 2015; Richardson et al., 2020). For example, status affects physical and mental health (e.g., Adler et al., 2000; Gray-Roncal et al., 2021; Reiss et al., 2019; Shahraki et al., 2018), employment (Morrison, 2019; Schuring et al., 2013), and access to education (e.g., Richardson et al., 2020). Importantly, social class conceptually and perceptually overlaps with the stereotype content and trustworthiness-dominance models. For example, people from lower socioeconomic backgrounds are frequently judged as less competent but warmer than their higher status counterparts (Carrier et al., 2014; Connor et al., 2021; Swencionis et al., 2017), suggesting a warmth-competence trade-off in social class stereotyping (Durante et al., 2017). Additionally, laypeople associate high social class standing with the experience of more positive valence (i.e., greater happiness; Diener & Biswas-Diener, 2002, 2009) which in turn relates to warmth and trustworthiness perceptions (Adams et al., 2012; Oosterhof & Todorov, 2009).

Such stereotypes exert an important influence on social perceptions, including from faces (e.g., Freeman et al., 2020; Kawakami et al., 2017). For example, recent research shows that when people consider two social traits to be conceptually similar, their impressions of those traits from faces also overlap (Stolier et al., 2018, 2020) and such conceptual and perceptual overlap is determined by group stereotypes (Xie et al., 2021). For example, participants base rich and poor judgments of others on facial features of class-related stereotypes such as warmth-stereotypes (Bjornsdottir & Rule, 2017). Yet, despite the centrality of social class in human societies and the clear link to central social trait dimensions, facial features that elicit social class perceptions remain poorly understood. As a result, it is unclear how perceptual overlaps between social trait and social class arise.

1.3 Drivers of social trait perception

Due to the centrality of social trait attributions, finding key features that elicit perceptions of these traits has been a central aim in the literature (Jaeger & Jones, 2022; Oosterhof & Todorov, 2008). Three potential sources of social trait impressions interact: Context, perceiver characteristics, and target characteristics. While all three play a role in social trait perception (Biancardi et al., 2017; Jaeger & Jones, 2022; Pace & Gnisci, 2019), the current work focuses on target characteristics and, specifically, face shape, face complexion, and facial expressions. However, this section gives an overview of each of the three factors – context, perceiver characteristics, and target characteristics.

1.3.1 Context

Most broadly, social judgments are shaped by the context in which they occur. Particularly, wealth (Keres & Chartier, 2016), threat (Brambilla et al., 2018; J. Wang et al., 2020), and social cues (Carragher et al., 2021; Hareli et al., 2018; Watkins et al., 2013) shape social trait perceptions. For example, placing faces in scenes conveying wealth increases their perceived trustworthiness (Keres & Chartier, 2016). Equally, while submissive faces are usually judged as more likely to help than dominant faces, this effect reverses in economic decisions (Hareli et al., 2018). Additionally, threatening contexts can change social perceptions. For example, priming women with angry male faces increases the salience of facial dominance cues (Watkins et al., 2013) and when workers face threats from their colleagues, such as freeloading, they prefer a more dominant looking leader (Bøggild & Laustsen, 2016). Similarly, trustworthiness perceptions are more important for voting decisions in peacetime than in wartime (Little et al., 2012), suggesting that prosocial traits carry higher value in positive, non-threatening situations. Finally, the social context also plays a role in trait perception. For example, people judge others differently if they encounter them in an online dating compared to a political campaign scenario (Todorov & Porter, 2014). Similarly, untrustworthy faces are judged as more trustworthy when they appear in a group (Carragher et al., 2021).

However, context is notoriously difficult to define and study as it spans a wide range of factors as illustrated above. Additionally, each added source of variance, such as context, non-linearly increases the complexity of social trait perceptions and their investigation (Jack & Schyns, 2017). It should also be noted that while context does shape social trait perceptions, social traits are encoded whether they are context-relevant (e.g., when buying a house) or not (Klapper et al., 2016). However, compatible contexts, such as viewing an untrustworthy face in a threatening context, do facilitate social trait perceptions (Brambilla et al., 2018; J. Wang et al., 2020). Overall, however, it is more tractable to initially focus on individual characteristics when studying social trait perception.

1.3.2 Perceiver characteristics

Social trait perceptions further vary depending on the perceiver's own characteristics. One major such factor is gender. For example, women generally assign higher trustworthiness ratings than men, particularly to female faces (Mattarozzi et al., 2015). However, experimental settings involving economic games lead to a reversal of this effect and greater trustworthiness ratings by men than women (Buchan et al., 2008; Chaudhuri & Gangadharan, 2007), indicating that context plays an important role in such gender differences. Another important factor are perceivers' personal attitudes and experiences. For example, racial bias influences trustworthiness perception with stronger racial biases leading to greater distrust in other-race faces (Stanley et al., 2011). Such racially biased social trait judgments stem from a difference in the visual processing of own and other-race faces (Charbonneau et al., 2020), underlining the importance of visual cues for social trait perception. Similarly, one's own personality influences social judgments. For example, people who score low on agreeableness and high on aggression rate unfamiliar faces as less trustworthy than their more agreeable and less aggressive counterparts (Mattarozzi et al., 2015). Likewise, dominant men assign lower dominance ratings to masculine looking faces than less dominant men, suggesting lowered sensitivity to facial dominance cues (Watkins et al., 2010). Even non-stable characteristics, such as lack of sleep, influence social trait perception (Alkozei et al., 2018). Overall, people also tend to prefer faces that look familiar (Dotsch et al., 2017) or bare familial resemblance (Bailenson et al., 2008; DeBruine, 2002; DeBruine et al., 2008).

However, impressions of some social traits are more heavily influenced by perceiver characteristics than others (Hehman et al., 2017). Specifically, perceiver characteristics play a larger role compared to target characteristics in the perception of competence but not in the perception of warmth or dominance (Hehman et al., 2017). It is therefore important to account for potential perceiver effects when investigating social trait perception. This can be done in a number of ways including 1) through tightly controlling participant characteristics, such as their gender and exposure to other cultures, 2) through modeling perceptions for each participant separately, thereby ensuring individual differences are preserved, or 3) through including a wide range of participants from different cultural, economic, and ethnic backgrounds. Here, I reduced perceiver influences primarily through the first two of these approaches – by modeling each individual participant's mental representations of different social traits while also ensuring minimal exposure to non-Western cultures and equal participant gender balance wherever possible.

1.3.3 Target characteristics

The final major predictor of social trait judgments are target characteristics. Besides the face, target characteristics that impact social trait perception include the voice (Lavan, Mileva, Burton, et al., 2021; Lavan, Mileva, & McGettigan, 2021; Ohala, 1982), clothing (e.g., Fleischmann et al., 2016), body pose (Biancardi et al., 2017; Rennung et al., 2016), gait (Satchell et al., 2021), and gestures (Gnisci & Pace, 2014; Pace & Gnisci, 2019). For example, lower voice pitch increases dominance ratings of both men (Ohala, 1982; Puts et al., 2006) and women (Borkowska & Pawlowski, 2011). Crucially, voices, much like faces, are perceived along two fundamental dimensions – trustworthiness and dominance (McAleer et al., 2014), underlining the wide-reaching importance of these key social traits. Similarly, social trait impressions from voices, like faces, happen quickly although at 400 ms (Mileva & Lavan, 2022) they are a little slower than those from faces. Additionally, the voice may play a greater role than the face in dominance perception but a smaller role than the face in trustworthiness perception (Rezlescu et al., 2015). Such dominance perceptions are linked to vocal masculinity (Wolff & Puts, 2010) akin to the link between facial masculinity and dominance perception (Sutherland et al., 2015). Indeed, people also make similar judgments about masculinity and femininity from faces and voices (Smith et al., 2016) and use both modalities (face and voice) to update initial impressions (Masi et al., 2021). Finally, humans perceive smiles from speech (Ponsot et al., 2018), though it unknown if here is a link between smile (i.e., happiness) perception and trustworthiness judgments as is observed for faces (Dotsch & Todorov, 2012; Knutson, 1996).

Another socially impactful target feature is body pose and, relatedly, gestures and gait (Maricchiolo et al., 2009; Thoresen et al., 2012). Specifically, expansive gestures and an overall space-consuming body posture increase ratings of dominance (Koppensteiner et al., 2016; Tiedens & Fragale, 2003) while decreasing ratings of trustworthiness (Koppensteiner et al., 2016). Similarly, a speaker appears more competent when using rhythmic gestures (Gnisci & Pace, 2014) linked to their speech or gestures directed at objects (Burgoon et al., 1990; Maricchiolo et al., 2009). In contrast, there is little evidence that body pose or gestures influence perceptions of warmth (Burgoon et al., 1990; Maricchiolo et al., 2009; Rennung et al., 2016). Overall, however, social judgments from faces appear to be more reliable than those from body posture (Rule et al., 2012) although this is likely to vary depending on the social context (Hostetter, 2011).

1.3.4 Face shape

The perhaps most widely studied social trait target cue is face shape. Face shape itself is a complex information space comprised of a multitude of features (Jack & Schyns, 2017; Jaeger & Jones, 2022) including, for example, overall face shape, eyes, nose, and mouth shapes. Because of this, researchers frequently focus on a reduced subset of these features. For example, babyfacedness – faces with a round shape, protruding forehead, large eyes, full cheeks, and a small nose and chin – received much attention in the impression literature (e.g., Berry & McArthur, 1985; Zebrowitz et al., 2003; Zebrowitz et al., 2011). Specifically, immature (i.e., babyfaced) looking faces elicit higher ratings of warmth, honesty, naïvety, and kindness than mature looking faces (Berry & McArthur, 1985; Jaeger & Jones, 2022; Oosterhof & Todorov, 2008). Perceived babyfacedness also correlates with other important social dimensions, such as perceived age (Berry & McArthur, 1985). However, babyfacedness is a relatively coarse measure which includes a number of facial features and it is not clear which of these in particular play a role in social trait perception. Additionally, babyfacedness is not a good indicator of true behavior (Poutvaara et al., 2009; Zebrowitz & Montepare, 2008). In an effort to focus on more behaviorally relevant features, researchers therefore turned to sexual dimorphism and in particular to Facial Width-to-Height Ratio (fWHR).

fWHR refers to the facial width, measured at the cheekbones, divided by the facial height, measured as the distance between the upper lip and mid-brow (Hehman et al., 2015). Higher facial masculinity, and with it fWHR, are linked to higher levels of testosterone and are thus seen as an 'honest signal' of testosterone levels, particularly in males (Carré et al., 2009). Indeed, facial masculinity and fWHR are believed to be associated with real-world behaviors such as aggression (Carré et al., 2010) and untrustworthiness (Arnocky et al., 2018; Stirrat & Perrett, 2010). Moreover, these features correlate with key social traits.

For example, both facial masculinity and higher fWHR correlate positively with perceived dominance and negatively with perceived trustworthiness and warmth (Albert et al., 2021; Burriss et al., 2007; Pivonkova et al., 2011; Stirrat & Perrett, 2010; Wen et al., 2020). Similarly, fWHR and its correlated features, including brow ridge height and chin depth, predict perceived aggression and threat (Carré et al., 2010). In addition to fWHR, other masculine features that drive dominance perceptions include the angle of the jaw and thickness of the lips (Burriss et al., 2007). Specifically, dominant and aggressive looking faces comprise a low brow, deep chin, a sharper jaw angle, thinner lips, and overall shorter, thinner faces (Burriss et al., 2007; Carré et al., 2010). Simply increasing how masculine a face appears increases dominance perception after exposures of as little as 100 milliseconds (Albert et al., 2021). However, recent evidence not only questions the link between perceived and actual dominance (H. Wang et al., 2019) but also suggests that sexual dimorphism, and indeed fWHR and babyfacedness, are far less relevant to social judgments than emotional facial content (Jaeger & Jones, 2022; D. Zhang et al., 2020).

Specifically, positive intent, i.e., trustworthiness and warmth, is associated with positive emotions, such as happiness (Hehman et al., 2015; Knutson, 1996; Oosterhof & Todorov, 2009). In contrast, ability, i.e., dominance and competence, is associated with negative emotions such as anger and disgust (Hehman et al., 2015; Knutson, 1996; Oosterhof & Todorov, 2009). Emotional facial content particularly drives perception of intent-related social dimensions whereas ability is more highly associated with facial width (Geniole et al., 2014; Hehman et al., 2015). Together, these findings demonstrate an important shortcoming of hypothesis driven approaches which focus on a priori defined facial features: While any given set of facial features may be correlated with social judgments, it may be missing other features crucial to social perception (Jack & Schyns, 2017).

In an effort to gain a more complete understanding of the face shape features associated with key social traits, Oosterhof and Todorov (2008) used reverse correlation, a data driven method (see subsection 1.4.1 for a discussion of this technique), to derive full face representations associated with trustworthiness and dominance perceptions. The advantage of this approach is that it simultaneously models *all* facial features of interest (e.g., all shape features) without making a priori assumptions about the importance of any given subset of features. However, this and similar later work faced new challenges. Firstly, Oosterhof and Todorov (2008) derived their two central traits, trustworthiness and dominance, based on dimension reduction of participants' free ratings of faces. However, as subsection 1.2.6 showed, there is substantial disagreement over the number and names of central social dimensions (Jenkins et al., 2011; A. L. Jones & Kramer, 2021; Sutherland et al., 2013). In particular, competence and dominance share only little facial information suggesting that they are separate, but both highly important, social traits (Oliveira et al., 2019). Additionally, Oosterhof and Todorov (2008) included only male stimuli. This was likely done to keep the experiment at a manageable length as reverse correlation experiments require high trial numbers (Jack & Schyns, 2017). However, substantial evidence shows that male and female faces elicit different social judgments (e.g., Albert et al., 2021; Burriss et al., 2007; Oh et al., 2019; Wen et al., 2020). Additionally, and perhaps due to this focus on male faces, Oosterhof and Todorov (2008) identified trustworthiness and dominance as orthogonal, i.e., decorrelated dimensions. In contrast, later work found a significant correlation between trustworthiness and dominance judgments for female faces (Sutherland et al., 2015). Finally, although reverse correlation models the entire face shape, the investigation of *which* of these shape features are particularly important to a given social trait judgment poses additional challenges.

To identify relevant features, extant work frequently relies either on visual inspection of group mean images (classification images) (e.g., A. L. Jones & Kramer, 2021; Oosterhof & Todorov, 2008; Sutherland et al., 2018; Sutherland et al., 2013), on comparisons of pixel luminance values (e.g., Dotsch & Todorov, 2012; Oliveira et al., 2019), or on judgments of classification images on related dimensions such as maturity or babyfacedness (Oosterhof & Todorov, 2008; Sutherland et al., 2013). For example, Sutherland et al. (2018) identified relevant features of their three social trait dimensions through visual inspection. However, such visual inspection cannot objectively identify which facial features drive social perceptions. To alleviate this limitation, a common approach is to obtain behavioral ratings of the classification images from a new set of perceivers and on a new set of potentially related dimensions. For example, Oosterhof and Todorov (2008) obtained ratings on relevant dimensions for each classification image, such as angry, happy, and facial maturity. This approach allows for a more objective identification of relevant facial features but nevertheless still relies on relatively coarse measures (e.g., happy).

In a more nuanced approach, Dotsch and Todorov (2012) applied clustering to the classification images' pixel luminance values to identify face regions that predict judgments of trustworthiness and dominance. This approach allows for the identification of specific face regions which elicit a given social judgment. Specifically, pixel luminance clusters helped identify the mouth, eye, and hair region as relevant to trustworthiness judgments and the eyes, hair, and chin angle as relevant to dominance judgments (Dotsch & Todorov, 2012). However, clustering of pixel luminance cannot identify what it is about a given facial feature that correlates with social judgments. For example, if clustering of pixel luminance indicates the eye region as relevant for trustworthiness judgments, it remains unclear whether the eyes have to be bigger, further apart, higher up or any combination of these to positively contribute to perceptions of trustworthiness. Secondly, this approach yields a poor resolution and only sporadic indicators of relevant features. For example, one pixel cluster identified by Dotsch and Todorov (2012) indicated only a small part of one eyebrow. Does this indicate that the eyebrows as a whole play a role in trustworthiness perception or is it evidence that social trait perception is lateralized and driven by highly specific types of facial features? Even once these issues are resolved, the common use of group level classification images still poses a substantial problem. Not only do classification images disregard any individual variation by averaging across all participants (Jack & Schyns, 2017) but they also increases type I error rates (Cone et al., 2020). As a result, despite the wide range of evidence for the importance of face shape in social perception, we do not have a clear understanding of the fundamental facial features that drive social perception.

1.3.5 Face complexion

In contrast to face shape, the literature investigating the role of face complexion, i.e., skin coloration, in social trait perception is sparse. This may in part be due to the greater impact of face shape compared to complexion on social perceptions (Oh et al., 2019). Additionally, investigating face complexion poses its own challenges. First, even small changes in the color display between monitors affect the perceived color (Thorstenson et al., 2018). Investigations of face complexion therefore require tight experimental control, usually in a laboratory setting (Thorstenson et al., 2018). Furthermore, although skin coloration differs between different regions of the face (Fink & Matts, 2008), skin color manipulations are frequently applied to the whole face due to technical limitations (Thorstenson et al., 2018; Thorstenson & Pazda, 2021). As a result, complexion is often overlooked in social trait perception and even studies that do include complexion variations frequently do not or cannot include the specific complexion features that drive social perceptions (Albohn & Adams, 2021; Dotsch & Todorov, 2012). Instead facial complexion is more frequently investigated in relation to age, health, and attractiveness (Fink & Matts, 2008; Fink et al., 2006). For example, 'patchier' skin color (as opposed to smooth skin color) increases perceived age, as do wrinkles and folds (Fink et al., 2012; Fink & Matts, 2008). However, these dimensions correlate with key social traits (Cuddy et al., 2005; Jaeger et al., 2018), making it likely that these complexion changes also impact social trait perception. And, indeed, when including skin coloration variations in the stimulus images, an additional youthfulness-attractiveness dimension emerges (Sutherland et al., 2018) which original investigations, with constant skin color, did not identify (Oosterhof & Todorov, 2008).

Despite these limitations there is clear evidence that face complexion plays a role in social trait perception (Oh et al., 2019). For example, facial redness increases perceived dominance and aggression but also friendliness and approachability (Stephen et al., 2012; Thorstenson & Pazda, 2021). Similarly, facial skin smoothness increases perceived trustworthiness (Tsankova & Kappas, 2016). Indeed, for competence, complexion may potentially play a more important perceptual role than face shape (Oh et al., 2019). In particular, darker skin tones and overall healthier looking faces elicit higher competence judgments (Sutherland et al., 2018) but also higher dominance judgments (Vernon et al., 2014). Similarly, face complexion plays a role in emotion perception (e.g., Thorstenson & Pazda, 2021), which itself is tightly linked with social trait perception (Hehman et al., 2017; Jaeger & Jones, 2022). Overall, humans appear to be particularly attuned to color differences on faces compared to non-faces (Thorstenson, 2018), suggesting that complexion does play a role in social perception.

1.3.6 Facial expressions

In addition to face shape and complexion, humans reliably perceive social traits from facial expressions (Gill et al., 2014; Hehman et al., 2017; Knutson, 1996). Research in this domain has focused almost exclusively on the perception of social traits from *emotional* facial expressions which subsection 1.2.2 discussed in more detail. For example, participants perceive angry facial expressions as dominant and happy facial expressions as trustworthy (Hareli et al., 2009; Knutson, 1996). For trustworthiness perception in particular, the mouth region plays an important role (Dotsch & Todorov, 2012), suggesting that it is particularly the smile resemblance which elicits trustworthiness and happiness judgments. Interestingly, however, smiles can also be perceived as dominant (Niedenthal et al., 2010; Tracy & Robins, 2008). Similarly, a smile does not always connote happiness but may convey a range of emotions including embarassement, amusement, or politeness (Ambadar et al., 2009; Keltner, 1995). It is therefore unclear whether it is indeed the happy facial expression resemblance that is crucial to trustworthiness perception or whether any type of smile would elicit high ratings of trustworthiness. As there are a wide range of possible smiles (Ambadar et al., 2009; Rychlowska et al., 2017), modeling the impact of each type of smile on social trait perception is technically challenging. Data-driven modeling of social trait facial expressions can therefore be a fruitful avenue in discerning the precise latent signaling structure that is driving similarities between social traits and emotional facial expressions (Jack & Schyns, 2017). Initial work in this area showed that dominance smiles are asymmetrical and include a raising of the upper lip while reward and affiliative smiles are more symmetrical (Rychlowska et al., 2017). Similarly, facial expressions that convey trustworthiness comprise raised eyebrows and a smile (Gill et al., 2014) which resembles the arched eyebrows and raised lip corners of neutral trustworthy faces (Dotsch & Todorov, 2012; Oliveira et al., 2019).

Most of the work investigating the link between emotional facial expressions and social trait perception has relied on static 2D facial expression images (Hehman et al., 2017; Knutson, 1996). However, facial expressions are inherently dynamic, 4D signals – comprising 3-dimensional space and time (Burt & Crewther, 2020). The time dimension may be critically important to the perception of, for example, emotions and may therefore similarly play a role in social trait perception (Burt & Crewther, 2020). Indeed, dynamic facial expressions can override first impressions made from face shape alone (Gill et al., 2014). For example, faces judged as dominant appear trustworthy when they display a trustworthy facial expression (Gill et al., 2014). There is also evidence that facial expressions include early, biologically rooted signals that may signal approach and avoidance before becoming more complex and discriminable into different emotions (Jack et al., 2014). Approach and avoidance signals may also play a fundamental role in social trait perception (e.g., A. L. Jones & Kramer, 2021), highlighting the potential nuance in the link between emotional expressions and social trait perception. Taken together this evidence suggests that (1) emotional facial expressions and social traits share a perceptual basis, (2) social traits are readily perceived from a range of dynamic facial expressions, and (3) both how dynamic facial expressions convey social traits and how these relate to emotional facial expressions is poorly understood.

1.3.7 Other facial features

Beyond face shape, complexion, and facial expressions, a range of other facial features are candidates for social trait perceptual influences. For example, head tilt and position are highly important for social trait perception and interact with eye gaze (Bee et al., 2010; Sutherland et al., 2017; Toscano et al., 2018). Specifically, an upward tilt of the head increases dominance (Mignault & Chaudhuri, 2003; Witkower & Tracy, 2019) and intimidation (Hehman et al., 2013). In contrast, a downward head tilt can either appear dominant if it displays observer-directed gaze (Hehman et al., 2013; Toscano et al., 2018; Witkower & Tracy, 2019) or submissive if the eyes are downcast (Mignault & Chaudhuri, 2003; Otta et al., 1994; Toscano et al., 2018). These effects may be related to facial width-to-height ratio which both raising and, to a lesser extent, lowering of the head increase (Hehman et al., 2013). Head position and gaze furthermore interact with emotional facial expressions. For example, happy facial expressions appear more trustworthy if the head is facing the perceiver rather than away from the perceiver (Sutherland et al., 2017). However, while eye gaze plays a role in social trait perception, there is as of yet no indication that eye color is similarly related to social trait perception (Kleisner et al., 2010; Kleisner et al., 2013).

Less well studied facial cues for social trait perception include head hair (Dotsch & Todorov, 2012; Macrae & Martin, 2007), facial hair (Bakmazian, 2014; Neave & Shields, 2008), and makeup (Klatt et al., 2016). For example, participants perceive women with lose hair as warmer than those with braided hair (Klatt et al., 2016) and women with brunette hair color appear more competent than those with red or blonde hair (Kyle & Mahler, 1996).

Similarly, bearded men appear more trustworthy (Bakmazian, 2014) and dominant (Neave & Shields, 2008) than men without beards. Finally, wearing makeup decreases perceived warmth (Klatt et al., 2016) but increases perceived competence (Aguinaldo & Peissig, 2021; Etcoff et al., 2011) and trustworthiness (Etcoff et al., 2011) in women. However, the effects of both hair and makeup on social trait judgments are likely to change with cultural shifts (Kyle & Mahler, 1996; Rosenthal, 2004). Taken together this evidence demonstrates that facial features that can be manipulated (e.g., makeup) play a little understood role in social trait perception. It is therefore imperative to control for these factors if one wishes to isolate specific face features of interest.

1.4 The current work

The previous sections raised a number of main questions that still remain within the social trait perception literature. Firstly, it is unclear what the fundamental, latent face features are that elicit perceptions of key social traits. Chapter 2 addresses this question and focuses on the latent face shape and complexion features that elicit perceptions of four key social trait dimensions – dominance, competence, trustworthiness, and warmth. Importantly, I model complexion features in the three color channels of the CIELAB (International Commission on Illumination L*a*b) color space: Lightness (dark to light), green-red, and blue-yellow. This color space is modeled after the human visual system, designed to be perceptually uniform, and therefore uniquely suited to studying human facial complexion (Thorstenson, 2018). Next, chapter 3 compares these social trait shape and complexion models to to the perception of another highly important and less well understood social dimension: social class. Finally, chapter 4 examines the dynamic facial expressions that elicit perceptions of these four social traits. It does so by mathematically modeling the precise facial movement patterns that elicit social trait judgments, represented as individual facial muscle groups or Facial Action Units (AUs) based on the Facial Action Coding System (Ekman & Friesen, 1978). This allows for precise control of the facial expressions and objective analysis of individual AU patterns. I then compare the resulting facial expression space and that of social trait face shape to emotional facial expressions, thereby locating the source of perceptual similarities between social traits and emotions and addressing fundamental questions about how social traits are inferred from features resembling dynamic facial expressions.

1.4.1 Methodological considerations

To derive these models of social trait and social class perception, I use reverse correlation. Reverse correlation is a psychophysical, data-driven technique that was first developed to study auditory signals (Ahumada & Lovell, 1971). However, it is now widely used in the social perception literature to examine what facial features are associated with a given social category, such as trustworthiness or dominance (e.g., Jack & Schyns, 2017; Oosterhof & Todorov, 2008; Sutherland et al., 2018). Reverse correlation relies on the random sampling of a given feature space, for example, of human faces, to derive a classification image corresponding to the participant's mental representation of a given social category. For example, in the most common approach, a base image of a face is repeatedly superimposed with random noise, thus altering the appearance of the underlying image. Participants then rate or categorize these new images. To create a classification image, researchers then average the superimposed noise across those trials that a participant rated as a given social category. In social perception literature in particular, reverse correlation has become increasingly common (Brinkman et al., 2017). It has successfully been used to model the facial appearance associated with key social traits (Dotsch & Todorov, 2012), and the similarity between the trustworthiness-dominance and warmth-competence models (Oliveira et al., 2019), to investigate how different social traits are visually integrated (Oliveira et al., 2019), and to explore cross-cultural differences in social trait perception (Sutherland et al., 2018). Reverse correlation is therefore an agnostic, data-driven approach with wide applications within social psychology.

Because of its data-driven nature, reverse correlation has several advantages for the study of human faces compared to traditional, hypothesis driven approaches (Brinkman et al., 2017; Jack & Schyns, 2017). Firstly, the face comprises a high number of features, each of which may or may not contribute to a given social judgment (Jack & Schyns, 2017). As a result, there are near infinite hypotheses a researcher may propose for which features elicit a given social attribution (Dotsch & Todorov, 2012). Even if a researcher successfully identified a specific facial feature using a hypothesis driven approach, they cannot be certain that they are not missing another important feature (Brinkman et al., 2017). Reverse correlation addresses this issue as it does not make a priori assumptions of which features are important and, theoretically, samples the entire information space of interest. Secondly, reverse correlation relies on participants' spontaneous use of the information presented (Brinkman et al., 2017). That is, participants utilize whatever cues they find most important to do the task. In contrast, hypothesis driven approaches frequently vary a dimension of interest (e.g., race) and then ask for a given social judgment (e.g., aggression). Here, participants have only one cue to base their decisions on – race (Brinkman et al., 2017). Finally, the reverse correlation method is generic and can therefore be used to investigate any measurable information space (Jack & Schyns, 2017), including here 3D face shape, 2D complexion, and dynamic facial expressions.

However, reverse correlation also has a number of potential disadvantages. Because each stimulus is random and may therefore not contain any diagnostic facial information, reverse correlation experiments typically require a high number of trials. This is particularly true for high-dimensional information spaces such as 3D shape and 2D complexion. A solution to this is to restrict the information space to a specific sub-dimension of interest to reduce the number of trials needed (Jack & Schyns, 2017). Additionally, the common technique of creating a group level classification image, thereby averaging across all participants, increases type I error rates (Cone et al., 2020). In this thesis, I therefore modeled social trait perception for each participant separately and made group-level inferences only from these participant-wise models. Overall, however, reverse correlation is a versatile and powerful tool to investigate social trait perception from faces.

Chapter 2

Social trait perception from faces is driven by a common feature space

2.1 Chapter abstract

Humans form rapid first impressions of others based on their facial appearance. Given the central importance of social perception to human societal functioning, a long-standing goal has been to identify the face features that drive these judgments. Here, I use a data-driven method to show that the perception of key social traits is driven by a set of latent face features. Using a 3D-face identity generator and a perception-based data-driven method, I mathematically model the 3D shape and 2D complexion face features that drive the perception of dominance, competence, trustworthiness, and warmth. Principal Component Analysis of the face feature models revealed a latent feature space that projects onto other socially relevant dimensions. These results have direct implications for existing theoretical models of social face perception and the design of socially interactive digital agents.

2.2 Introduction

Humans readily infer the social traits of others, such as competence and warmth (Oosterhof & Todorov, 2008; Sutherland et al., 2018), from their facial appearance. Such rapid social attributions (Willis & Todorov, 2006) influence social interactions (De Neys et al., 2017; DeSteno et al., 2012; Re & Rule, 2016) and decision making (Franklin & Zebrowitz, 2016; Joo et al., 2015; Lyons & Simeonov, 2016; Valentine et al., 2014; Wilson & Rule, 2015). With such wide-reaching implications for human social life, a central goal has been to identify the main dimensions along which such social judgments are made (Fiske et al., 2002; Oosterhof & Todorov, 2008) and to reveal the fundamental facial features that drive these social perceptions (e.g., Jaeger & Jones, 2022). Influential models propose that social trait judgments are made along two main dimensions – dominance/competence (ability) and trustworthiness/warmth (intent) (Fiske et al., 2002; Oosterhof & Todorov, 2008). While such models map the conceptual relationships between different social traits (Stolier et al., 2018), it remains poorly understood what facial features drive these social perceptions.

The list of facial features that have at one point or another been thought to be relevant to social perception is long, ranging from facial Width-to-Height ratio to babyfacedness (Valentine et al., 2014; Zebrowitz et al., 2011). For example, perceived happiness and babyfacedness (e.g., round face, high forehead, large eyes, small nose and chin) of a face decrease perception of dominance and competence but increase perception of trustworthiness and warmth (Knutson, 1996; Montepare & Dobish, 2003; Oosterhof & Todorov, 2008). Yet, it is unclear which feature of happy expressions, for example the corners of the mouth or dimples, drives this effect. Similarly, babyfacedness encompasses a range of shape features (Berry & McArthur, 1985), each of which may or may not be contributing to social trait perception. Even less is understood about the contribution of complexion information to social trait perception. With the exception of competence, complexion information contributes less than shape to the perception of social traits (Oh et al., 2019). However, it is again unclear which specific complexion features, such as redness of the cheeks or contrasts around the eyes, contribute to these perceptual differences.

Moreover, the fact that all four social traits correlate along perceptual dimensions such as babyfacedness and emotionality (Knutson, 1996; Montepare & Dobish, 2003; Oosterhof & Todorov, 2008) suggests that judgments of these four social traits emerge from a latent face feature space. Yet, while there is a drive to find those face features which are fundamental to social perception (Jaeger & Jones, 2022), such investigations are empirically challenging due to the complexity of the human face (e.g., Jack & Schyns, 2017) which comprises both 3D shape and 2D complexion features. It therefore remains an open question which specific facial features are shared across central social traits and are thus fundamental to social perception. This chapter addresses this question using a data-driven method based on subjective social perception to mathematically model the specific face features that drive the perception of four key social trait dimensions – dominance, competence, trustworthiness, and warmth – in each of 30 participants.

2.3 Data-driven modeling of face features associated with perceptions of key social traits

Given that the face is a highly complex visual stimulus, identifying precisely which face features drive particular social perceptions is empirically challenging. To do so, I used a data-driven approach that agnostically samples naturalistic variations in face features and tests them against subjective human social perception (Zhan et al., 2019). Subsection 2.3.1 through subsection 2.3.7 and Figure 2.1 illustrate this method.

2.3.1 Participants

Thirty white Western participants completed the experiment (15 female, 15 male; mean age = 23.77 years, Standard Deviation (*SD*) = 5.03 years). To reduce the potential impact of cross-cultural differences in social trait perception (Sutherland et al., 2018), participants completed a questionnaire assessing exposure to non-Western cultures (see Supplementary Material – Participant questionnaire). All participants had normal or corrected to normal vision with no symptoms of synesthesia, psychological, psychiatric or neurological conditions that can affect face-processing, as per self-report. Participants gave written informed consent and were paid £6/hour for participation. The experiment was approved by the University of Glasgow College of Science and Engineering Ethics Committee.

2.3.2 Stimuli

To randomly generate faces on each experimental trial, I used a Generative Model of 3D Face Identity (GMF) that is based on high resolution 3D captures (14,319 3D vertex coordinates, 1024×536 RGB pixels) of real people (402 total, 232 female, 170 male; mean age = 28.19 years, *SD* = 14.65 years; 245 white, 149 East Asian, 8 Black) and has a high fidelity generative capacity (for full details see Zhan et al., 2019). Specifically, the generative model represents variance in 3D shape and 2D complexion according to (1) sex, ethnicity, age, and the interactions between them, derived using a General Linear Model (GLM) applied to the vertex coordinates and pixel values of the faces in the database, and (2) individual face identity, derived using a Principal Components Analysis (PCA) applied to the vertex coordinate and pixel value residuals. Thus, identity variance is represented as 402 PCs for 3D shape and 402 PCs (for each of five spatial frequency bands) for 2D complexion.

To randomly generate a face identity, I added to the age, ethnicity, and sex GLM averages (i.e., 18-35 years, white, female or male), the 402 identity components with randomly sampled PC scores for 3D shape and 2D complexion (see Figure 2.1A left). I generated a total of 2,400 such face identities (1,200 female, 1,200 male). Next, I applied to each face identity a 3D texture layer (representing fine details such as pores and wrinkles) randomly selected from an individual face in the database of the same ethnicity and gender. Finally, I fitted each face identity to a standard 3D head and neck according to the size and shape of each individual face identity.

2.3.3 Experimental task

Each participant rated the same 2,400 face stimuli on each of the four social trait dimensions in separate tasks for a total of 9,600 trials ([1,200 3D faces $\times$ 2 stimulus sex $\times$ 4 social trait dimensions]; see section 6.2 in the Supplementary Material for an exploration of how many trials are needed to derive stable social trait face models). Participants completed the experiment across 12-13 separate sessions of 1 hour each with no more than 3 sessions per day and at least a 1-hour break between each two sessions. Within each session, participants completed 5 blocks of 160 trials each for a total of 800 trials per session with short breaks after each block. In each block of 160 trials, the social trait dimension task and stimulus sex remained constant. Each participant therefore completed a total of 15 blocks per social trait dimension. I randomized the order of blocks and the stimuli across the experiment for each participant. At the start of each block, I informed participants of the sex of the stimulus faces in the current block and the target social trait.

On each trial, participants viewed the randomly generated 3D face on the left side of the screen but within the participant's center of vision and rated it on a vertically arranged scale ranging from 1 (e.g., very submissive) to 7 (e.g., very dominant), with 'don't know' as the central fourth button. Both the stimulus and response options remained on screen until response. A white fixation cross appeared during the 500 ms inter-stimulus-interval following response. Participants responded using a Graphical User Interface (GUI). I displayed all face stimuli on a 1920 $\times$ 1080 resolution color-calibrated flat panel monitor at a constant viewing distance of 72 cm, thereby subtending 14.25° (vertical) and 8.74° (horizontal) of visual angle, reflecting the average size of human faces (e.g., Ibrahimagić-Šeper et al., 2006) during typical social interaction distances (e.g., Hall, 1966). I ensured a constant viewing distance using a chin rest and controlled the experiment using Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007) in MATLAB R2018a.

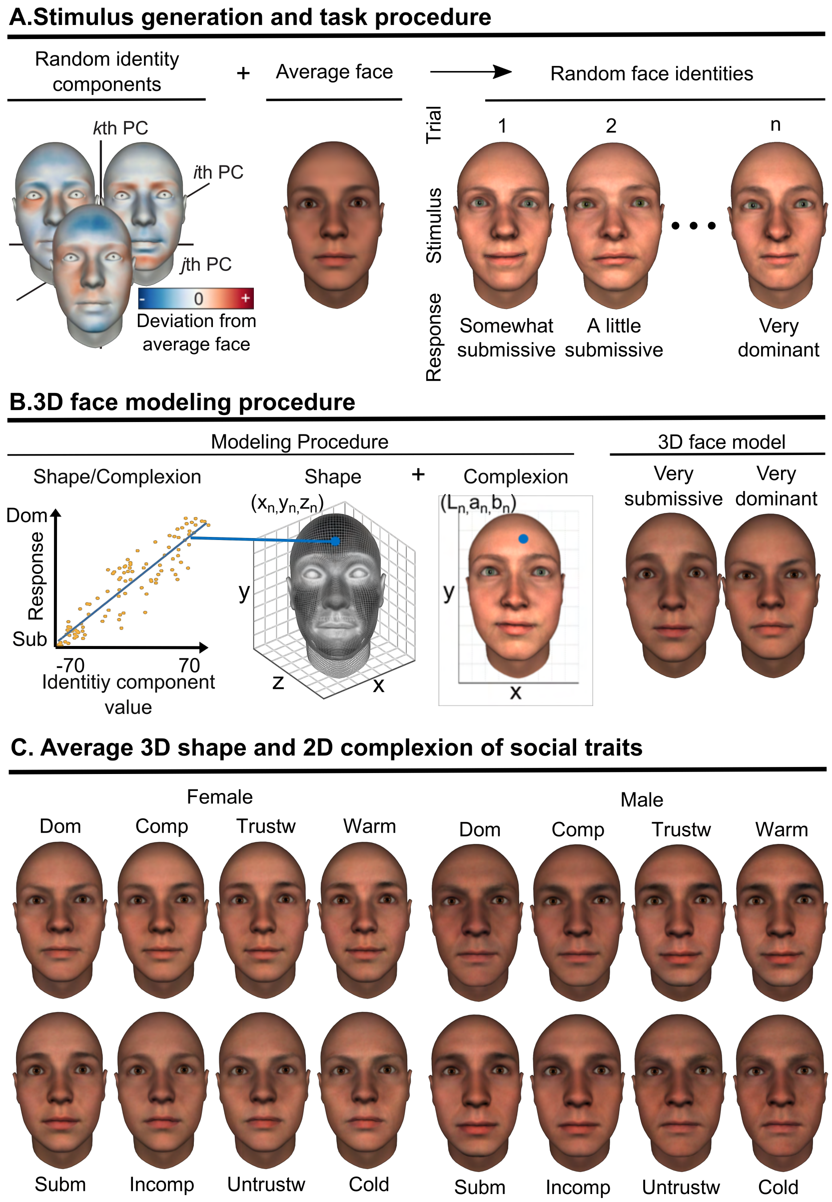

Figure 2.1. Modeling the 3D face features that drive social trait perception. (A) Stimulus generation and task procedure. On each trial, a generative model of human face identity generated a novel face stimulus by randomly sampling naturalistic face feature variations. Participants rated each face stimulus on each one of four social traits on a 7-point bipolar scale (e.g., from 'very submissive' to 'very dominant'). (B) 3D face modeling procedure. To model the face features driving each social trait perception, I measured the statistical relationship between the randomly generated face features presented on each trial and the participant's responses. This produced a statistical model of the face features that drive the participant's social trait perceptions (see example for one participant on right). (C) Average 3D shape and 2D complexion of social traits (median across participants), extrapolated for clarity.

2.3.4 Social trait face modeling procedure

Next, to model the specific face features that elicit social trait judgments, I measured the statistical relationship between the 3D face shape and 2D complexion features presented on each trial and the participant's responses using linear regression (see Figure 2.1B left for an illustration; see also section 6.3 in the Supplementary Material for an exploration of the appropriateness of the linearity assumption to model these data). Specifically, for each social trait dimension (dominant-submissive, competent-incompetent, trustworthy-untrustworthy, warm-cold), I mass-univariately estimated the slope and intercept for each 3D shape coordinate (x, y, z), each 2D complexion pixel in CIELAB color space, each social trait dimension, stimulus sex, and participant separately using ridge regression. This yielded a mathematical model describing, separately, the shape and complexion information for each social trait dimension for a total of 240 social trait models per shape and complexion [2 stimulus sex $\times$ 4 social traits $\times$ 30 participants].

This agnostic data-driven approach produces statistically robust quantitative models of the face features that drive social trait perceptions that can then be analyzed to objectively specify these face features. Deriving individual participant models also enables the preservation and representation of individual variance from which population variance can be estimated (Ince et al., 2021). Figure 2.1B right shows an example resulting face model for the perception of dominant-submissive for one participant. I then validated the resulting 240 models prior to further analysis.

2.3.5 Validating 3D social trait face shapes

To validate each social trait 3D face shape model, I recruited a new set of 40 participants (white, Western, 20 female, 20 male; mean age = 23.21 years, *SD* = 4.27 years), using the same eligibility criteria as described above (see subsection 2.3.1). Using the 240 face shape models [30 participants $\times$ 4 social trait dimensions $\times$ 2 sex of face], I generated from each model six faces ranging from e.g., 'very submissive' to 'very dominant' including extrapolations and excluding the 'don't know' midpoint, resulting in a total of 1,440 faces [6 generated faces $\times$ 240 models; 720 females, 720 males]. I fixed face complexion by using the average model. I then created all pairwise combinations of faces from each individual model (excluding identical pairs) for each sex separately, resulting in 3,600 pairs of faces [30 participant face models $\times$ 4 social traits $\times$ 2 stimulus face sexes $\times$ 15 pairings of the 6 faces]. Each individual face pair appeared once for each social trait scale end (e.g., "choose which face looks most submissive", "choose which face looks most dominant"), resulting in 7,200 total unique trials. Participants viewed on each experimental trial, a pair of same-sex faces generated from the same model, presented side-by-side (randomized across participants), and judged which of the two appeared most like a given social trait (e.g., "choose which face looks most submissive") in a 2-alternative forced choice task. Participants responded by clicking on one of the faces, with the next stimulus pair appearing after an inter-stimulus-interval of 500 ms. I instructed participants to work quickly and base their judgments on their first impressions. Each participant completed a total of eight separate rating tasks – i.e., most submissive, dominant, trustworthy, untrustworthy, competent, incompetent, warm, cold – conducted in separate blocks.

Each participant completed a pseudo-random subset of 3,600 of the total 7,200 face pair trial [15 participant face models $\times$ 8 tasks $\times$ 2 stimulus face sexes $\times$ 15 pairings of the 6 faces]. Trials were split into 16 blocks of 225 trials each, with each block comprising one judgment task (e.g., most dominant) and the same sex of face stimuli. I randomized the order of the stimuli in each block and the block order across the experiment for each participant and provided short breaks after each block. I displayed all face stimuli in the same way as described above (see subsection 2.3.1). Prior to testing, all participants provided informed consent and completed 10 practice trials. Participants completed the experiment across between four and five 1-hour-long sessions and took at least a 1-hour break after each set of four blocks.

Following the validation experiment, I computed for each of the 240 individual participant social trait shape models and for each validation participant separately, the number of times the participant chose the correct face in the pair, i.e., recognition accuracies. For each social trait shape model separately, I then applied a one-sample *t*-test at alpha = .05 to these recognition accuracies [20 validation participants $\times$ 6 faces per social trait, e.g., from very submissive to very dominant = 120 recognition accuracy values per model] to test whether participants choices reflect above change accuracy (i.e., above mean = 50%). After Bonferroni-Holm correction for multiple comparisons, the majority (95.0%; 114 female, 114 male; 93.33% competent, 98.33% warm, 100% dominant, 88.33% trustworthy) of individual participant social trait models performed with above chance accuracy. I excluded 12 models (5.0%) that performed at below chance accuracy from subsequent analyses (see Figure 6.10A in the Supplementary Material for individual model recognition accuracies).

2.3.6 Validating 2D social trait face complexions

To validate the 2D social trait complexion models, I performed a Leave-One-Out Cross Validation (LOOCV). A computational validation approach ensured that the relatively larger effects of face shape compared to complexion did not mask any complexion-specific effects during validation. Further, LOOCV most closely approximated the shape validation procedure as I proceeded in the following steps:

1. For each social trait, sex of stimulus face, and participant separately, I mass-univariately

regressed the responses of 29 of the 30 participants against the stimulus complexion information, represented as the 402 identity residual PCs [402 PCs $\times$ 5 spatial frequency bands] using ridge regression. Next, to identify identity residual PCs that are significantly associated with response, I applied Monte Carlo simulation (two-tailed) with 5,000 iterations.

2. I then trained a Generalized Linear Mixed Effects Model (GLME) on the 29 participants' responses and only significant (see step 2) identity residual PCs. Specifically, the GLME included fixed effects for each significant PC as well as random effects of participants.

3. Finally, I used the model obtained in step 3 to predict response of the remaining participant. I then correlated (Spearman's ρ) predicted with true responses to assess how well the model predicted responses of the left-out participant.

I repeated steps one through three for each of the 30 participants, leaving out a different participant each time. I thus obtained a correlation value corresponding to how well each individual participant's 2D complexion model was predicted by the other participants' data for each social trait and sex of stimulus face. I applied Bonferroni-Holm corrections to the *p*-values of these correlations and considered complexion models with $p \leq .05$ to be validated. Using this method, a total of 220 complexion models were validated (91.7%; 108 female, 112 male; 85% competent, 95% warm, 95% dominant, 91.67% trustworthy; see also Figure 6.10B in the Supplementary Material for individual model performance). In sum, using LOOCV conceptually most closely approximated the shape validation as it tested each participant's perception of each social trait against another set of participants' perceptions of the same social trait. Figure 2.1C shows the validated 3D face models (shape and complexion), averaged (median) across participants, for each social trait dimension and sex of face separately.

2.3.7 Visualizing the face features of the social trait face models

Following validation of the face models, I analyzed them to formally characterize and visualize the specific face features of each. To do so, I first derived significant face features using permutation testing. Specifically, I performed a non-parametric permutation analysis (two-tailed) with 1,000 iterations on the 3D shape (x, y, z separately) and 2D complexion (L*a*b separately) of each validated social trait model. Next, I compared these face features to the average face (i.e., white female or male, 18–35 years). Specifically, I derived predicted 3D shape vertex values (x, y, and z coordinates) and 2D complexion pixel values in L*a*b based on each individual validated social trait model (3D shape: n = 228 ; 2D complexion: n = 220). I then computed the difference between these predicted values and those representing the average face from the generative face model. For 3D shape, I subtracted the average face's x, y, and z vertex values from those of each predicted face and calculated the cosine of the angle between the model's difference from average (i.e., residual) and the vector vertical to the tangent of each vertex of the average face. This produced a single value per 3D face vertex, describing the magnitude and direction of the difference between the predicted model vertex and the average face. For 2D complexion, I used a similar approach, calculating the difference from the average face, for L (luminance), a (green-red), and b (blue-yellow) separately. Figure 2.2 shows the resulting validated and significant face features (see also Supplementary Material Figure 6.11 for results separated into x, y, and z plane and Figure 6.12 for replication of results across participants for each social trait).

2.4 Results

Figure 2.2 shows the median differences to the average face for each social trait and sex of stimulus face separately. For shape, blue shows the face features that deviate inward compared to the average face; red shows the face features that deviate outward compared to the average face (see colorbar on right; see also Figure 6.11 in the Supplementary Material for results separated into x, y, and z planes). For example, male faces perceived as dominant, competent, untrustworthy, and cold are each characterized by protruding eyebrows and chins as shown by the red regions. In contrast, female faces perceived as warm and trustworthy are characterized by upturned corners of the mouth and raised eyebrows as shown by the blue regions. For complexion, results are presented for each color channel separately – Lightness, green-red, blue-yellow. Here, coloration shows complexion features that differ from the average face – for example, lighter/redder/yellower complexion (see Supplementary Material – Figure 6.13A for green-red results in colorblind friendly colors). For example, male faces perceived as dominant, competent, untrustworthy, and cold each comprise darker, greener, and bluer (i.e., cooler) complexions with lighter eyebrows compared to the average face. Similarly, female faces perceived as dominant, incompetent, untrustworthy, and cold each comprise darker and bluer complexion with lighter, yellower eyebrows compared to the average face. A visual inspection of these features across social traits suggests that the perception of certain social traits, such as dominant and competent or trustworthy and warm, are driven by similar face features, as is suggested by the correlated behavioral judgments of such faces (Oosterhof & Todorov, 2008; Walker & Vetter, 2016, see also Figure 6.14 in the Supplementary Material).

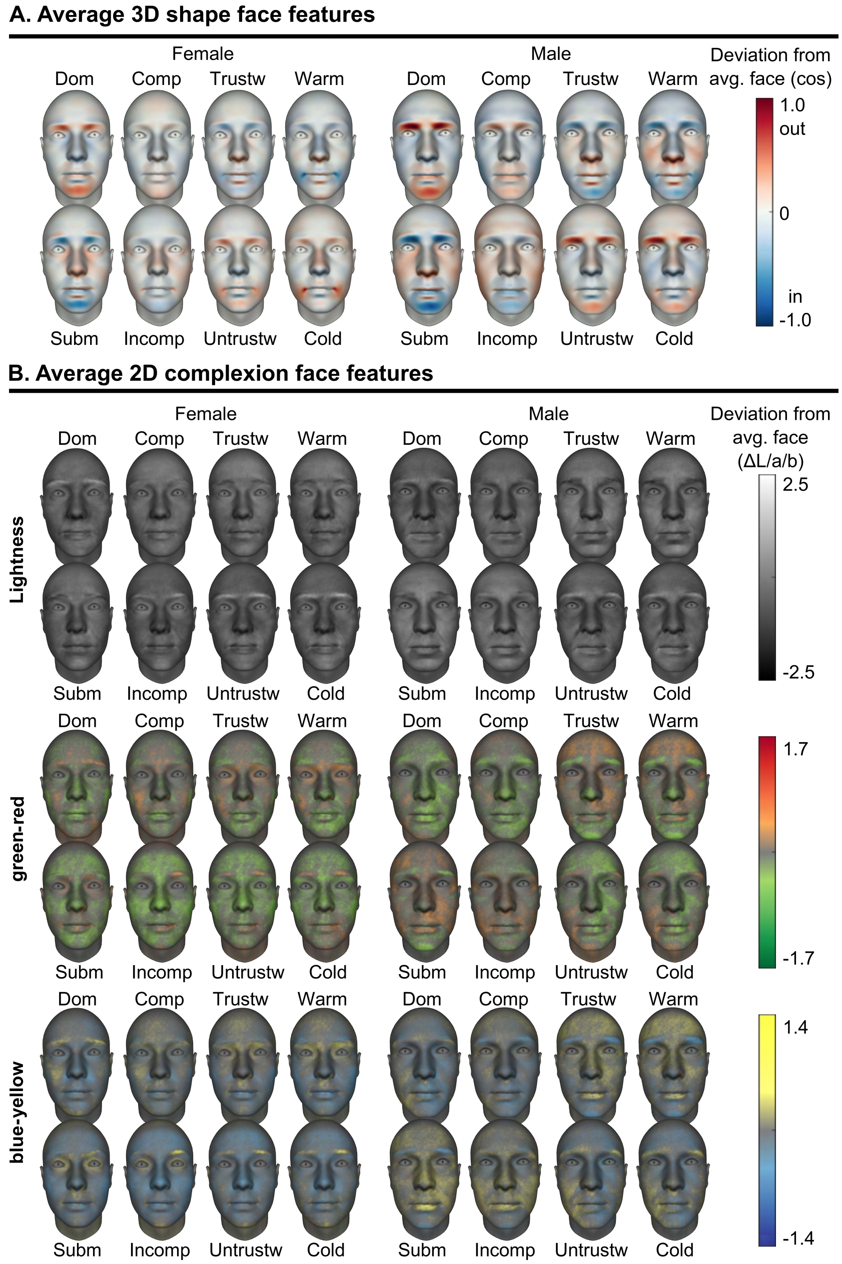

Figure 2.2. Average 3D shape and 2D complexion of social traits. (A) Average 3D shape face features. Face maps show the shape variations relative to the average face (normalized). Red indicates outward features relative to the average; blue indicates inward features (see colorbar). For example, dominance is associated with protruding, lowered eyebrows/chin, and flatter cheek bones. (B) Average 2D complexion face features per color channel (L*a*b), social trait, and sex of stimulus face (normalized within each color channel). Coloration indicates absolute differences in complexion compared to the average face. For example, dominance has an overall darker complexion with lighter eyebrows.

2.4.1 Modeling the latent face features of social trait perception

To identify the latent face feature space underlying social trait judgments, I applied a Principal Component Analysis to the set of validated social trait face models for each sex of stimulus face and for shape and complexion separately. Specifically, for shape, I first transformed the significant and validated shape face information into a 42,957 [14,319 3D shape vertices $\times$ 3 x, y, z coordinates] by 228 [validated trait models = 4 social traits $\times$ 2 trait poles (e.g., dominant and submissive) $\times$ 30 participants – 12 non-validated models] matrix for each sex of stimulus face separately. Similarly, I transformed the down-sampled complexion information into a 183,654 [342 $\times$ 179 pixels $\times$ 3 L*a*b color channels] by 220 [validated trait models = 4 traits $\times$ 2 trait poles (e.g., dominant and submissive) $\times$ 30 participants – 20 non-validated models] matrix for each sex of stimulus face. I then applied a PCA to each of these shape and complexion matrices separately, thereby obtaining PCA coefficients for each 3D shape vertex and each 2D complexion pixel. Finally, I triangulated the optimal number of PCs using a combination of the elbow method, second derivative, and Mutual Information analysis (see Supplementary Material – Determining optimal social trait PCs and Figure 6.6 to Figure 6.8 for details). For each significant PC, I then reconstructed the full 3D face shape and 2D complexion – for L*a*b separately – by multiplying the 42,957 shape PCA coefficients and the 183,654 complexion PCA coefficients by the minimum and maximum PCA score for each PC separately.

2.4.2 3D shape

For 3D face shape, results revealed two main feature spaces for female faces (65.44% of variance explained) and two for male faces (71.37% of variance explained). Figure 2.3A shows these two feature spaces for each sex of face separately, each displayed as color-coded face maps on different axes (PC1 – y-axis; PC2 – x-axis). Colored face regions show the difference in 3D shape from the average male and female face, respectively (see colorbar to right; see also Supplementary Material – Figure 6.15 and Figure 6.16 for results per x, y, z plane separately for easier identification of relevant features). Specifically, for both male and female faces, the first feature space (PC1, represented in Figure 2.3A as the y-axis) comprises variations in protruding/raised eyebrows, flatter/protruding cheek bones, and a larger/smaller chin. In contrast, the second feature space (PC2, represented in Figure 2.3A as the x-axis) differs across male and female faces. For male faces, the features comprise variations in narrower/wider jaw, a protruding/flatter mouth and upturned/downturned mouth corners. For female faces, the features primarily comprise variations in downturned/upturned corners of the mouth, a larger/smaller chin, and flatter/protruding nose bridge.

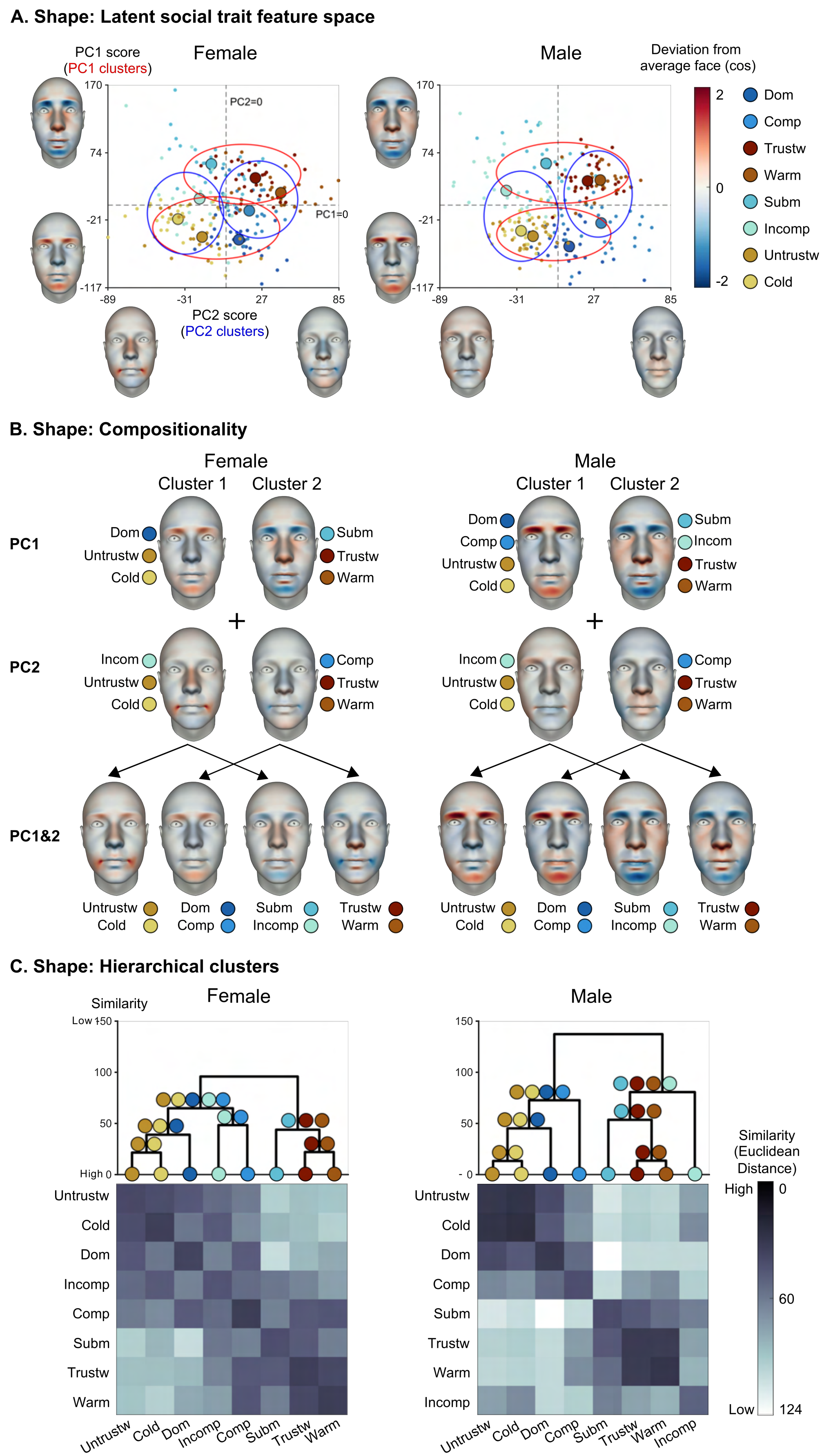
A. Shape: Latent social trait feature space
PC1 score
(PC1 clusters)
Female
Male
Deviation from
average face (cos)
170
74
-21
-117
-89
-31
27
85
PC2=0
PC1=0
PC2 score
(PC2 clusters)
2
0
-2
Dom
Comp
Trustw
Warm
Subm
Incomp
Untrustw
Cold
B. Shape: Compositionality
Female
Male
Cluster 1
Cluster 2
PC1
PC2
PC1&2
Dom
Untrustw
Cold
Subm
Trustw
Warm
Comp
Incom
Incomp
C. Shape: Hierarchical clusters
Similarity
Low 150
100
50
High 0
Female
Male
Untrustw
Cold
Dom
Incomp
Comp
Subm
Trustw
Warm
Similarity
(Euclidean
Distance)
High
Low
0
60
124

Figure 2.3. Latent face shape feature space subtending key social trait perceptions. (A) Latent feature space. PCA applied to the 3D face shape models revealed two main feature spaces, each represented here as color-coded face maps on separate axes (PC1 – y-axis, PC2 – x-axis). Red shows outward features compared to the average; blue show inward features (see colorbar on right). Scatterplots show the distribution of individual participant face models (small points) according to their PC scores and color-coded by social trait (see key to right). Larger points represent the median across participants and dashed lines correspond to a score of zero on each PC. Colored ellipses indicate clusters of each PC in red (PC1) and blue (PC2) based on K-means analysis. Together these show that PC1 creates two main clusters which are further subdivided into four main groups by PC2. (B) Compositionality. To illustrate how the two PCs compose individual trait subgroups and the face features associated with each, the top row shows trait clusters (K-means) with associated facial features based on the first PC. The second row shows trait clusters with associated facial features based on the second PC. Below, I show the subclusters of social traits with associated facial features formed based on both PCs. (C) Hierarchical clustering. Dendrogram plots and similarity matrices below show the average similarity of these face features across social traits, computed using pairwise Euclidean distance between PC scores (PC1 and PC2) of all models (darker tones indicate higher similarity; lighter tones indicate lower similarity – see colorbar to right). For example, in male faces, untrustworthy (light brown) and cold (yellow) comprise highly similar face features.

Having derived the main feature spaces that represent these social trait face models, I then examined how the individual participant models of each social trait distribute within these feature spaces. Specifically, I first assessed which social traits each PC was significantly associated with using correlations (Pearson's r) between PC scores and binary social trait vectors (e.g., 'dominant' and 'not dominant'). Most comparisons, except for dominant/submissive (PC2) and female competent/incompetent (PC1), were significant ($p \leq .05$ after Bonferroni-Holm correction; see also Table 6.1 and Table 6.2 in the Supplementary Material). Next, for each PC and sex of face separately, I applied K-means clustering to the social trait model PC scores, including only social traits which were significantly associated with a given PC. In all cases, a 2-factor solution best represented the data as assessed via silhouette scores (see Determining optimal K-means clustering solution in the Supplementary Material for details). Figure 2.3A shows the results as a scatter plot where each small point represents an individual participant model, color-coded by social trait (see legend on right). Larger points represent the average (median) across participants per social trait. Dashed lines mark zero PC values for each PC. Clusters of PC1 are indicated as red ellipses and of PC2 as blue ellipses, plotted at 1.5 *SD* around the centroid for each cluster (see also Supplementary Material – Determining optimal K-means clustering solution).

As shown by the distribution of the face models and clusters within each two-dimensional feature space, the social trait models form two broad clusters on each PC which, together, form four main sub-groups – submissive and incompetent (top left quadrant), trustworthy and warm (top right quadrant), and untrustworthy and cold (bottom left quadrant), and dominant and competent (bottom right quadrant). For example, submissive and incompetent face models cluster in the upper left quadrant, indicating that they are characterized by arched eyebrows, shorter chins, downturned mouth corners, and wider faces. To further explore the division of social traits based on these two feature sets, Figure 2.3B shows the face features associated with each social trait cluster based on the each PC only and then based on both PCs combined. I classified social traits as belonging to a cluster if more than 12.5-16.57% of models in this cluster were of this social trait ($\geq$1/number of traits with significant correlations; see Supplementary Material – Figure 6.17 for cluster proportions per social trait). The results demonstrate that, the first feature set clusters social traits into two main groups and these clusters are then further subdivided by the second feature set. For example, dominant, competent, untrustworthy, and cold cluster together based on protruding eyebrows and chin (PC1). However, untrustworthy and cold are characterized by downturned mouth corners and wider faces than dominant and competent (PC2). Additionally, the clusters reveal the facial features that drive perceptual differences between dominance and competence (Oliveira et al., 2019). Specifically, dominance is more strongly characterized by lowered eyebrows and a thinner mouth (PC1) while competence is more strongly characterized by upward arching corners of the mouth (male and female PC2) and narrower face width (male PC2). Indeed, for female faces competence was only significantly associated with the latter features (mouth corners), indicating a clear difference in competence perception from male and female faces.

Finally, to confirm the presence of these clusters based on both PCs, I computed the pairwise similarities between the PC scores of all 3D face models (i.e., the Euclidean distance between each pair of points in each feature space) for each sex of face separately. Figure 2.3C shows the results, averaged across models for each social trait, displayed as a gray-scale matrix where darker tones indicate higher similarity between social traits and lighter tones indicate lower similarity (see colorbar to right). Each matrix is ordered according to the hierarchical similarities of the average social trait PC score with the associated dendrogram displayed above each matrix with color-coded points representing each social trait. Results confirm that the social traits cluster hierarchically according to shared face features. For example, for male faces, the face models of untrustworthy, cold, dominant, and competent form a main cluster based on arched eyebrows, protruding cheeks, and longer, protruding chins (first feature space – PC1) with two sub-clusters (untrustworthy and cold; dominant and competent) based on face width and mouth shape (second feature space – PC2). In contrast, for female faces, the face models of competence are similar to those of submissiveness, trustworthiness, and warmth based on the upturned corners of the mouth (second feature space – PC2). In sum, dominant, competent, untrustworthy, and cold share a common feature set, specifically, the first feature space. These four traits are then further distinguished based on the second feature space.

2.4.3 2D complexion

For complexion, the PCA revealed three main feature spaces for female faces (38.67% of variance explained) and three main feature spaces for male faces (39.74% of variance explained). As for shape, Figure 2.4A shows these three feature spaces for each sex of face separately, each displayed as color-coded face maps on different axes (PC1 – y-axis; PC2 – x-axis, PC3 – z-axis). Colored regions show the difference in 2D complexion from the average male and average female face, respectively and for each color channel (see colorbars to right; see also Supplementary Material – Figure 6.13B for green-red results in colorblind friendly colors). Specifically, for both male and female faces, the first feature set (PC1, represented in Figure 2.4A as the y-axis) comprises on one hand lighter, redder, and yellower than average eyebrows and dark folds around the mouth (see in Figure 2.4A high scores) and on the other, dark, green, and blue contrasting eyebrows and light areas around the mouth (see in Figure 2.4A low scores). In contrast, the second feature set (PC2, represented in Figure 2.4A as the x-axis) affects overall face coloration with high scores corresponding to lighter, redder, and yellower face complexion and low scores to darker, greener, and bluer complexion. Finally, the third feature set (PC3, represented in Figure 2.4A as the z-axis), comprised on one hand darker, greener, and bluer forehead and eye regions but also lighter, redder, and yellower mouth regions and light folds around the mouth (see in Figure 2.4A high scores). On the other hand, it comprises lighter, redder, and yellower forehead and eye regions with darker, greener, and bluer mouth regions and dark folds around the mouth (see in Figure 2.4A low scores).

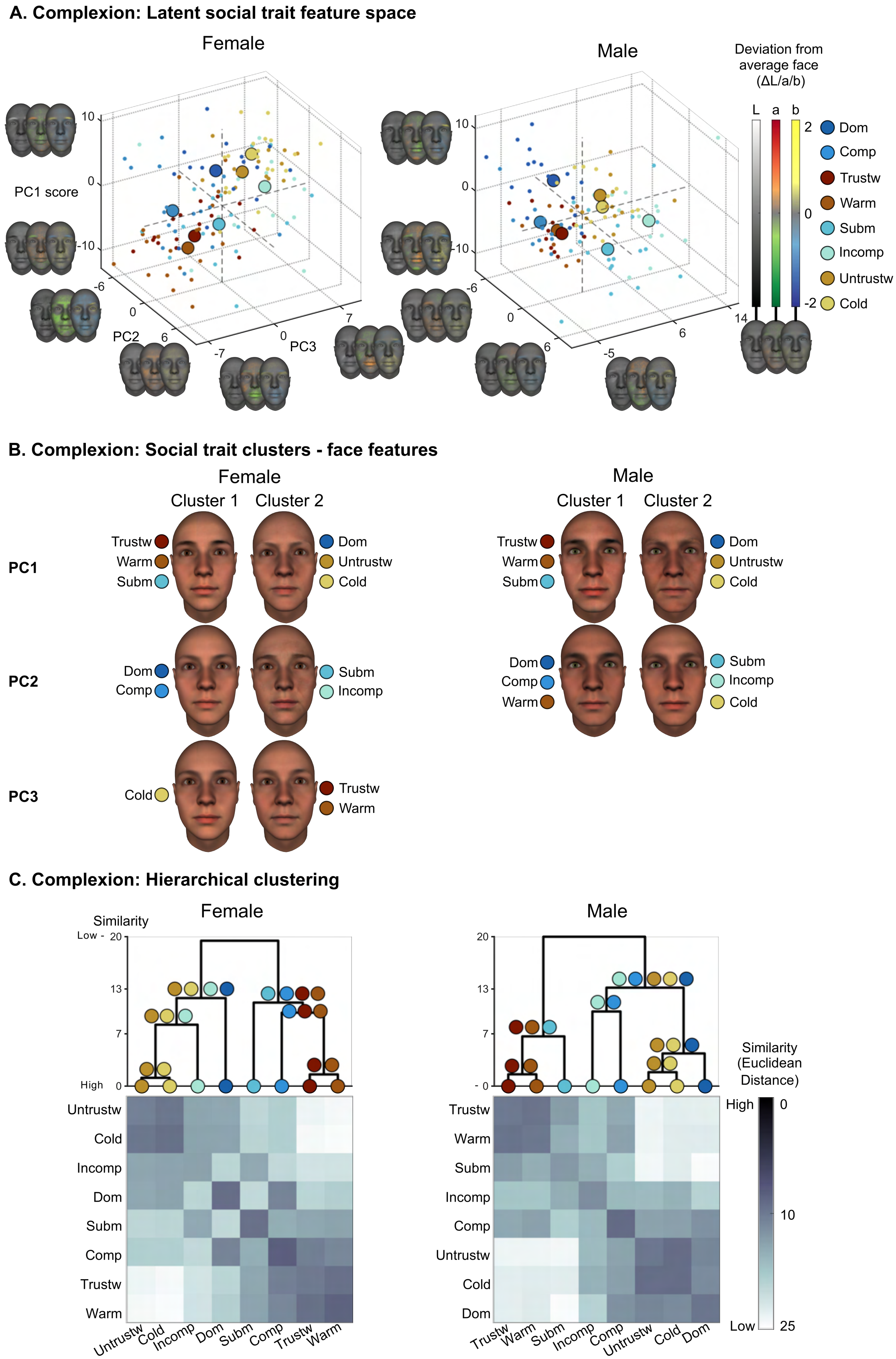
A. Complexion: Latent social trait feature space
Female
Male
Deviation from average face (ΔL/a/b)
L a b
2
0
-2
Dom
Comp
Trustw
Warm
Subm
Incomp
Untrustw
Cold
PC1 score
PC2
PC3
B. Complexion: Social trait clusters - face features
Female
Male
Cluster 1
Cluster 2
PC1
PC2
PC3
C. Complexion: Hierarchical clustering
Similarity
Low
High
Similarity (Euclidean Distance)
Untrustw
Cold
Incomp
Dom
Subm
Comp
Trustw
Warm

Figure 2.4. Latent face complexion feature space subtending key social trait perceptions. (A) Latent social trait feature space. To test for a common feature space, I repeated the analysis steps of the shape analysis (PCA) which yielded three PCs (PC1 – y-axis, PC2 – x-axis, PC3 – z-axis). Colored face maps show the complexion features of each PC compared to the average face for each color channel – Lightness (L), green-red (a), blue-yellow (b). Scatterplots show the distribution of individual participant face models (small points) according to their PC scores with larger point size indicating median scores across participants (points above 1.5 *SD* excluded for visualization purposes; see Figure 6.18 in Supplementary Material for full figure). Dashed lines indicate zero PC values for each PC. (B) Social trait clusters. Complexion face features (as texture face maps) of each sub-cluster of each PC as derived via K-means analysis (see Figure 6.20 in Supplementary Material for results in L*a*b separately). For each cluster, labels indicate the social traits captured by the cluster. (C) Hierarchical clustering. Clustering of the eight social traits according to their PC scores (all three PCs). Specifically, dendrogram plots showing how dissimilar (height of lines) any two given social traits are (median Euclidean Distance across participant-wise comparisons). Colored points here refer to each social trait. For example, in male faces, untrustworthy (light brown) and cold (yellow) are highly similar. These relationships are further illustrated by correlation matrices with dark colors indicating high similarity (e.g., between untrustworthy and cold in male faces) and light colors high dissimilarity (e.g., untrustworthy and submissive).

Next, I examined the similarity of these features across social traits by projecting each individual 2D face model into the three-dimensional feature space according to their PC scores and evaluating their distributions. Figure 2.4B shows the results as a scatter plot where each small point represents an individual participant model, color-coded by social trait (see legend on right). Note that Figure 2.4A excludes data points beyond 1.5 *SD* for clearer visualization of the results due to the overall smaller complexion effects and more tightly clustered social trait models (see Figure 6.18 for full plot). Larger points represent the average (median) across participants per social trait. Dashed lines mark zero PC values for each PC. As shown by the distribution of the face models in Figure 2.4A within each three-dimensional feature space, the social trait models cluster into two main sub-groups – dominant, incompetent, untrustworthy and cold, and submissive, competent, trustworthy, and warm. For example, dominant, incompetent, untrustworthy and cold scores cluster in the upper half of the plot indicating that these traits are characterized by light, non-contrasting eyebrows and darker, more contrasting, folds around the mouth.

To test these similarities further, I applied K-means clustering to these PC scores. However, I examined social trait clusters within each PC separately as correlations between PC scores and each social trait (as for shape above) showed that not all PCs were significantly associated with all social traits and indeed for male PC3 none of the social traits were significantly correlated (see Supplementary Material – Table 6.3 and Table 6.4 for descriptive statistics and correlation results per social trait and PC). Specifically, as for shape, for each PC and sex of face separately, I applied K-means clustering to the social trait model PC scores, including only significant social traits ($p \leq .05$; Pearson's r, corrected with Bonferroni-Holm). In all cases, a 2-factor solution best represented the data as assessed via silhouette scores (see Determining optimal K-means clustering solution and Figure 6.19 in the Supplementary Material for details). Figure 2.4B shows the resulting social trait clusters and their corresponding complexion face features as 2D texture face maps (see Figure 6.20 in Supplementary Material for results in L*a*b separately). This revealed that trustworthy, warm, and submissive share overall lighter complexions with darker, contrasting eyebrows, and redder cheeks (PC1, cluster 1) than average. In contrast, dominant, untrustworthy, and cold share lighter eyebrows, darker shading around the corners of the mouth and, particularly for male faces, darker folders around the mouth (PC1, cluster 2). However, for PC2 and PC3 female and male complexion features diverged. For female faces, submissive and incompetent (PC2, cluster 2) exhibited overall darker and more uneven complexion than their opposite trait counterparts. Additionally, for female but not male faces, trustworthy and warm were characterized by darker folds around the mouth (PC3, cluster 2).

Finally, to test for overall social trait clusters based on all three PCs, I computed the pairwise similarities between the PC scores of all 2D face models (i.e., the Euclidean distance between each pair of points in each feature space) for each sex of face separately. Figure 2.4C shows the results, averaged across models in each social trait, as a gray-scale matrix where darker tones indicate higher similarity between social traits and lighter tones indicate lower similarity (see colorbar to right). Each matrix is ordered according to the hierarchical similarities of the social traits, as determined by a dendrogram analysis displayed above each matrix with color-coded points representing each social trait. Results confirm that the social traits cluster systematically on the basis of shared complexion features. For example, for male faces, the face models of trustworthy, warm, submissive, and incompetent form a main cluster with two sub-clusters (trustworthy and warm; submissive and incompetent). In contrast, in female faces trustworthy and warm are clustered closely with competent but less so with submissive. Similarly, competent, untrustworthy, cold, and dominant are clustered closely in male faces while in female faces, dominant is more closely clustered with incompetent.

2.4.4 Social traits and related social judgments

The previous sections showed that the perception of key social traits is driven by a latent structure of two face shape feature spaces and three complexion features spaces that together subdivide social traits into four subgroups. Notably, for both shape and complexion, several of the identified face features such as upturned mouth corners (shape) and dark folds around the mouth (complexion) are also implicated in other social perceptions such as emotion (e.g., Jack et al., 2016) and age perception (e.g., Tsankova & Kappas, 2016). The latter result is striking as, in extant work, youthfulness emerged as a potential third central social trait (Sutherland et al., 2018; Sutherland et al., 2013; Vernon et al., 2014), suggesting that age-cues play a central role in social trait perception. I therefore next tested how each feature space identified in previous sections correlated with (a) affective signals (for shape only) and (b) age cues.

To test the relationship between central social trait shape and complexion features and emotion-related facial expression feattures and age cues, I correlated their respective 3D face shape and 2D complexion features across the whole face. Specifically, for shape, I first identified Action Units (i.e., individual facial movement patterns) that were highly relevant to the perception of at least one of the six basic emotions using data from previous experiments (happy, surprised, fearful, disgusted, angry, sad; assessed via population prevalence; Ince et al., 2021; Jack et al., 2014). Modeling the relationship with individual Action Units as opposed to full emotion models allowed me to identify the specific features that correlate between emotions and social traits as opposed to overall feature similarities. Next, I created 3D shape representations for each emotion-related AU (n = 38) separately using a Generative Face Grammar (GFG; Yu et al., 2012, see subsection 4.3.2 for details). Similarly, using the GMF, I created 3D shapes corresponding to the average youngest and oldest face shape (18 and 35 years, corresponding to the age range of stimuli in the reverse correlation experiment), keeping all other variables (sex and ethnicity) constant. Finally, for each social trait shape PC, I reconstructed the 3D shape based on both low and high PC scores. Next, for social trait PC, AU and age shape models separately, I normalized (z-score) their values for each dimension (x, y, z) separately. To reduce noise, I then identified shape vertices that varied highly for each social trait PC (>70% of variance). Finally, I correlated (Pearson's r) the resulting 3D shape vectors of each social trait PC with each AU and age shape model. I proceeded similarly for face complexion for age only (no data was available for emotion complexions). Figure 2.5 shows the results.

Specifically, Figure 2.5A shows the correlations between the face features of each PC (PC1 – y-axis; PC2 – x-axis) and the face features of each AU (positive correlations only). Several individual AUs are labeled for illustration. Colors indicate comparisons of female faces (green) and male faces (yellow) and equally colored ellipses illustrate the distribution of correlation values within the two-dimensional social trait PC space. As shown by these ellipses, for male faces, primarily face features of PC1 (see colored face maps) correlated with AUs. In particular, higher scores on this PC1, corresponding to arched eyebrows,

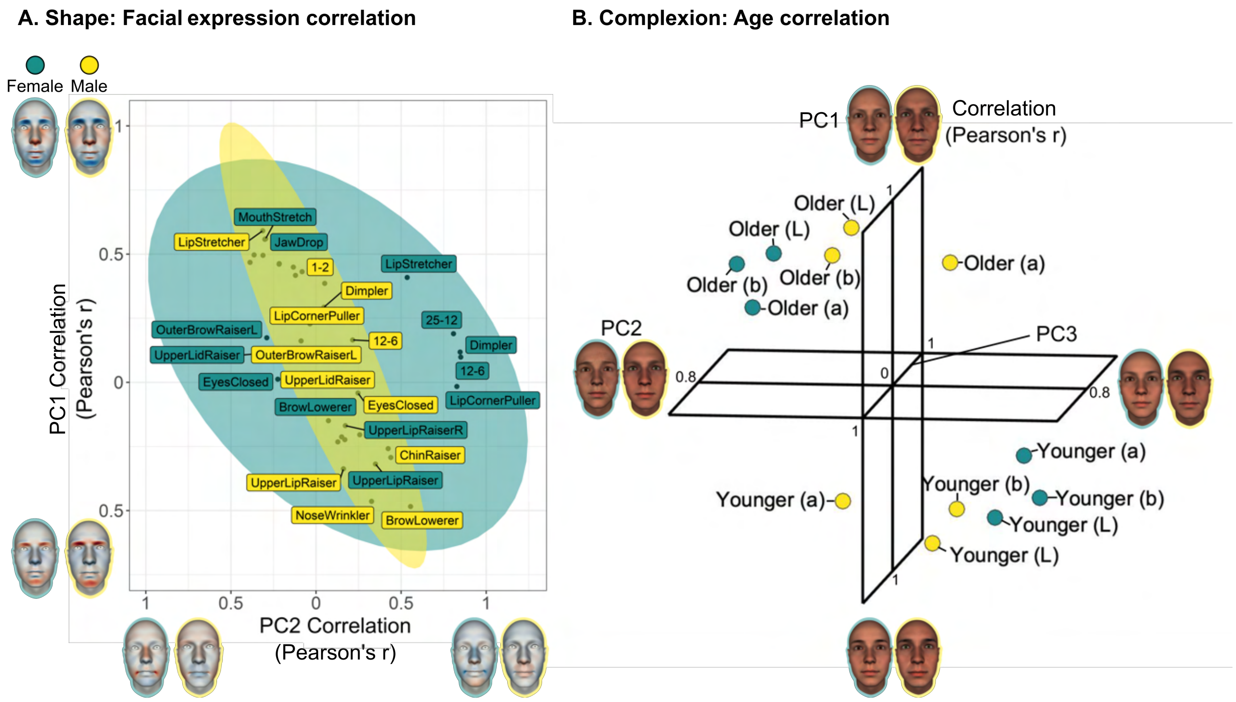

Figure 2.5. Correlations between latent social trait face features and relevant social judgments. (A) Shape: Facial expression correlation. A scatterplot shows the correlation (Pearson's *r*) between each social trait shape PC (PC1 – y-axis; PC2 – x-axis) and each one of 38 emotion-relevant AUs. Ellipses indicate the distribution of these correlation values in the two-dimensional space of female (green) and male (yellow) comparisons. Labels indicate the names of several relevant AUs. Negative correlations are excluded for clearer visualization. Colored face maps correspond to social trait PC features as in Figure 2.3. (B) Complexion: Age correlation. In a 3D scatterplot, colored points (female = green; male = yellow) show positive correlations between complexion values of younger (average face at 18 years) and older (average face at 35 years) faces in a similar manner as for shape and for L*a*b separately. Labels indicate the corresponding age and L*a*b color channel. Faces correspond to social trait PC features for PC1 (horizontal) and PC2 (vertical) as in Figure 2.4. Face features of PC3 (diagonal) are excluded for visualization purposes as only female face trustworthy, warm, and cold significantly correlated with this PC (see Table 6.3).

puffier cheeks, and longer noses and chins, correlated positively with Lip Stretcher, Mouth Stretch, Jaw Drop and Eyebrow Raiser. In contrast, lower scores, corresponding to lowered, protruding eyebrows, shorter noses and shorter chins, correlated positively with Nose Wrinkler, Upper Lip Raiser, and Brow Lowerer, which are common in negatively valenced facial expressions (e.g., Jack et al., 2016). For female faces, both PC1 and PC2 correlated with AUs due to correlations between PC2 (high scores, corresponding to raised corners of the lips and narrower faces) and AUs such as Dimpler, Lip Corner Puller, and Lip Stretcher which are all part of happy facial expressions (e.g., Jack et al., 2016).

Similarly, Figure 2.5B shows the positive correlations between each social trait complexion PC and younger and older face complexion in L*a*b separately (female = green; male = yellow). Note that social trait shape correlated less strongly with age-related cues and these results are included in the Supplementary Material (Figure 6.21) only. For both female and male faces, age complexion correlated with PC1 such that PC1 features forming the dominant/untrustworthy/cold cluster (slightly darker complexion with darker folds around the mouth and less contrasting eyebrows) positively correlated with older age. In contrast, PC1 features of the submissive/trustworthy/warm cluster (lighter complexion with dark contrasting eyebrows) correlated positively with younger age. Similarly, PC2 correlated with age such that the incompetent/submissive cluster (darker complexion) correlated with older face complexion.

2.4.5 Relative contributions of shape and complexion

The previous section suggests that shape and complexion play different roles in social trait perception. I therefore next tested the relative contribution of shape and complexion information to social trait judgments using a 10-fold cross-validation. Specifically, at each iteration I trained a simple linear model on the face information (402 identity components of shape, complexion, or shape and complexion combined) of a sub-sample of trials (n = 1,080 of 1,200 trials) to predict participants' social trait ratings (i.e, 'very dominant' to 'very submissive'). I did this for each sex of stimulus face, social trait, and participant separately. I then tested each model on the holdout data and finally correlated (Spearman's ρ) the predicted responses obtained at each iteration with the true responses. Figure 2.6 shows the resulting per-participant correlation values as colored points for models trained on shape (red), complexion (blue), and shape and complexion combined (purple) with summary statistics represented as boxplots.

For all social traits and both female and male faces, the models trained on shape and complexion jointly outperformed both of the individual models. Additionally, for most social traits, including dominance, trustworthiness, and warmth, the shape model outperformed the complexion model. However, for competence (both sexes), the model performed as well

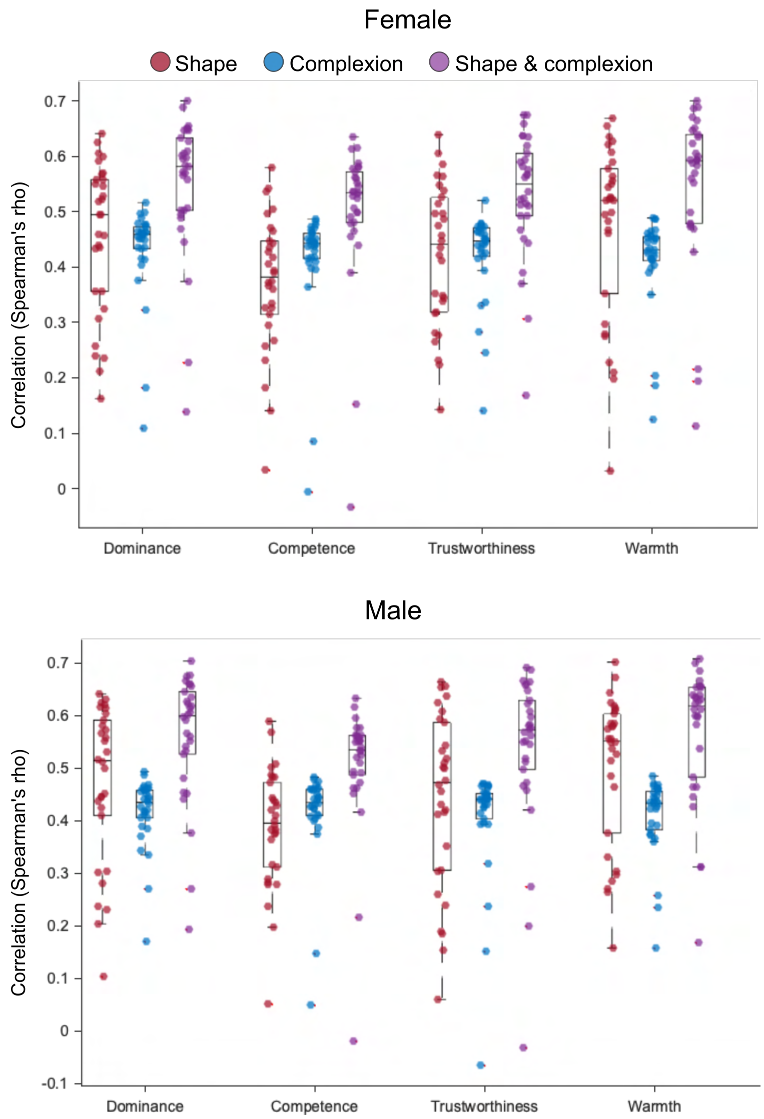

Figure 2.6. Social trait shape and complexion feature contribution. Boxplots and individual model points show the correlation (Spearman's ρ; y-axis) between predicted and observed values of a 10-fold cross-validation procedure. At each iteration, I trained a linear model based on the stimulus identity components of shape only (red), complexion only (blue), and shape and complexion combined (purple), for each sex of stimulus face and each social trait dimension separately.

as or better when trained on only the complexion face information compared to the shape information. However, the combination of shape and complexion still outperformed either of the individual models.

Taken together, these results show that the perception of social traits is driven primarily by a latent face feature space of 3D shape but also a latent 2D complexion feature space. These feature spaces are furthermore systematically associated with facial features of other socially relevant dimensions such as emotional expressions and age. Additionally, there are some differences between male and female faces in the perception of some of these social traits and how they relate to other social dimensions (e.g., positive affect).

2.5 Discussion

Chapter 2 addressed the longstanding search for latent face features driving social trait perception and showed that perceptions of key social traits are structured by two 3D shape and three 2D complexion face feature sets. Specifically, I used reverse correlation to model the shared 3D shape and 2D complexion features of four key social trait dimensions – dominance, competence, trustworthiness, and warmth. The results reveled that social traits systematically cluster according to two shape feature sets and three complexion feature sets. For example, trustworthy and warm are highly similar on the basis of narrower face width, arching eyebrows, protruding cheeks, upturned mouth corners, and darker, contrasting eyebrows. I furthermore showed that these central features correlated with emotion-related facial movement patterns for 3D shape (PC1 for male faces, PC1 and PC2 for female faces) and with age cues for 2D complexion (PC1 and PC2). Finally, I showed that for all social traits except competence, face shape best predicted social judgments but shape and complexion combined outperformed shape alone.

The two central 3D shape feature sets identified here together cluster social traits into four subgroups: dominant-competent, submissive-incompetent, trustworthy-warm, and untrustworthy-cold. These correspond closely to models of social trait perception, namely the trustworthiness-dominance model and stereotype content model that suggest that social trait perception is structured along two orthogonal dimensions often summarized as intent (dominance and competence) and ability (trustworthiness and warmth) (Fiske et al., 2002; Fiske et al., 2007; Oliveira et al., 2019; Oosterhof & Todorov, 2008). Such commonalities suggest that social trait perception could be driven by a latent set of face features that project onto two dimensions – intent and ability. However, while the results presented in this chapter support the idea of two orthogonal social trait dimensions of intent and ability, they also reveal that the latent shape features driving the perception of these dimensions do not directly equate to 'ability features' and 'intent features' but rather combine together to form the four subgroups of intent and ability.

Additionally, I showed that social traits and emotional facial expressions share face shape features such that in particular the first PC correlated with Action Units of negatively valenced facial expressions such as disgust and anger. These emotions in turn relate to avoidance (e.g., Curtis et al., 2011; Stins et al., 2011), aligning these results closely with recent findings suggesting that approach/avoidance is a main dimension of social trait judgments (A. L. Jones & Kramer, 2021). Not only does this support the hypothesis that the relationship between emotion and social trait perception can be characterized by approach and avoidance (Knutson, 1996) but it also contextualizes recent findings that emotion-similarity is highly predictive of trait judgments (Jaeger & Jones, 2022) by identifying the precise relationship between emotional expressions and social trait perception. These findings are also in line with the emotion overgeneralization theory which posits that positively valenced social traits such as trustworthy and warm are perceived as such because they resemble positive emotions such as happy (Montepare & Dobish, 2003; Said et al., 2009). Moreover, I show that emotion-related features play a larger role in social trait perception in female than in male faces, in particular for competence perception. These results correspond to the western stereotype that women are more emotional than men (Brescoll, 2016; Shields et al., 2017), but also to true physical differences between male and female faces where female faces physically resemble emotions such as happy and fearful (Hess et al., 2009).

These results furthermore resolve two conflicting findings in the literature. Firstly, recent evidence suggests that the face information that drives perceptions of the two ability-related traits, competence and dominance, diverges (Oliveira et al., 2019) and I show here that this is due to the importance of face-width in male faces and lip corner height in female faces for the perception of competence but not dominance. Secondly, while some work has identified two fundamental dimensions of social perception (A. L. Jones & Kramer, 2021; Oosterhof & Todorov, 2008) and other work three (Sutherland et al., 2018; Sutherland et al., 2013; Vernon et al., 2014), I show here that the perception of the third social dimension, age, is driven primarily by complexion rather than shape face features. In particular, the first complexion feature set which includes contrasts in the eyebrow region and wrinkles around the mouth was associated with age-related variance.

Such wrinkles and folds are more common in old age and increase perception of age while decreasing attractiveness perception (Aznar-Casanova et al., 2010; Samson et al., 2010). Similarly, stark color contrast drives perception of health and youth (Russell et al., 2016). Accordingly, previous work has shown that older looking faces are perceived as less warm and less healthy than younger faces (Zebrowitz et al., 2003). Such perceptions of age are largely driven by complexion and texture cues with only minor contributions by shape (O'Neil & Webster, 2011). However, skin complexion is rarely studied in relation to social trait perception even though it has been shown that overall face complexion makes a significant contribution to social trait perception (A. L. Jones et al., 2012; Oh et al., 2019). For example, perceptions of dominance are driven by darker face complexion (Vernon et al., 2014) which I found here is a reflection of overall darker face complexion in negative traits such as dominant, untrustworthy and cold. Overall, however, social judgments were better predicted by shape than by complexion face information, with the exception of competence, replicating earlier work (Oh et al., 2019), although shape and complexion together outperformed shape alone. I therefore show, for the first time, the specific latent complexion features that drive the perception of key social traits and which are closely linked to age-related complexion features.

In conclusion, chapter 2 showed that social trait perceptions are driven by a latent set of shape and complexion face features. Together, these dimensions form a feature space which clusters social traits into ability and intent subgroups. These results support the fundamental assumption that trait perception is primarily structured along two dimensions, namely ability and intent but show, for the first time, the latent face features underlying these perceptions and the specific contributions shape and complexion face features make to social trait perception.

Chapter 3

Social class perception is related but not equal to social trait perception

3.1 Chapter abstract

Social class is a central hierarchical structure in human societies that significantly impacts individual lives and wider societal functioning. Social class standing (e.g., whether someone is rich or poor) is, like other social attributes, judged from the face. It remains unknown what facial features drive the perception of social class and how these relate to other central social dimensions, specifically social traits. This chapter addresses this question by objectively modeling the 3D facial features that drive the perception of social class in 30 Western participants using a data-driven approach analogous to that of chapter 2. Analysis of the resulting 3D face models showed that while some facial features are common across social class and social traits, no one social trait is isomorphic to social class. The results reveal the specific facial features that drive the perception of social class in a Western culture, showing that the perception of social class is related but not equal to other core social judgments.

3.2 Introduction

Chapter 2 identified the latent facial features that subtend social trait perception. The centrality of these facial features in social perception suggests that they may hold importance for other key social concepts. One such concept is social class. In human societies, social class is a central social hierarchy (e.g., Kraus et al., 2013) that affects health (e.g., Adler et al., 2000; Sapolsky, 2004), access to education, and employment (Bjornsdottir & Rule, 2017; Kraus et al., 2019; Morrison, 2019; Richardson et al., 2020; Rivera, 2012; Rivera & Tilcsik, 2016; Whitty, 2001) and inclusion in economic opportunities (Nelissen & Meijers, 2011). Importantly, social class *conceptually* relates to social traits. For example, individuals of high social class standing are stereotyped to be more competent (Durante et al., 2017; Spencer & Castano, 2007), warm (Fiske et al., 2002), and to have experiences of more positive valence (i.e., greater happiness; Diener & Biswas-Diener, 2002) than individuals of low social class standing.

Such conceptual overlaps between social class and social traits suggest that these judgments may also share a *perceptual* overlap. Such top-down influences of social information (e.g., stereotypes) on social perception is demonstrated in extant work showing that when people consider two social traits to be conceptually similar, their impressions of those social traits from faces also overlap (Stolier et al., 2018, 2020). Additionally, such conceptual and perceptual overlap between social traits is determined by group stereotypes (Xie et al., 2021). Correspondingly, social class and social traits exhibit some perceptual overlap – faces judged as higher in social class are also perceived as more competent, warmer, and more positively valenced (Bjornsdottir & Rule, 2017, 2020; Bjornsdottir, 2019).

However, although, like social traits, social class is often judged from facial appearance (Bjornsdottir & Rule, 2017), the specific facial features that drive these critical social class judgments and the extent of their intrinsic relationship to social trait judgments remains unknown. According to general accounts of communication (Bradbury & Vehrencamp, 2011; Edelman & Gally, 2001; Shannon, 1948; Tinbergen & Tinbergen, 1948) and human perceptual expertise (Schyns et al., 1998; Tanaka & Taylor, 1991) the face, as a complex and rich source of information, could elicit correlated perceptions – for example, rich and competent – based on the same or different facial features. Identifying these facial features and examining their relationships to different social perceptions thus has important implications for understanding human social perception. This chapter examines this critical missing link, between facial features of social class and social trait perception, by using methods analogous to those used in chapter 2, namely reverse correlation, to model the face shape and complexion eliciting social class perceptions and comparing them to those of social traits.

3.3 Data-driven modeling of face features associated with perceptions of social class

3.3.1 Participants

Thirty white, English-speaking British participants (15 females, 15 males; mean age = 22.30 years, *SD* = 3.69 years) took part in the experiment. All participants completed a pre-screening questionnaire to assess whether they had lived in the UK most of their lives, that both of their parents were born in the UK, that they were of middle socioeconomic status (scoring 4-7 on the 1-10 MacArthur scale of subjective social status; Adler et al., 2000; *M* = 5.45, *SD* = 1.19), and that they had minimal experience with and exposure to non-Western cultures (e.g., de Leersnyder et al., 2011, see Supplementary Material – Participant questionnaire). All participants self-reported normal or corrected-to-normal vision, and no symptoms of synesthesia, or of psychological, psychiatric, or neurological conditions that could affect face processing. Each participant provided written informed consent prior to testing and received a standard rate of £6/hour for their participation. The College of Science and Engineering at the University of Glasgow provided ethical approval.

3.3.2 Stimuli

This experiment used the same stimuli as chapter 2 and as illustrated in Figure 3.1A. Specifically, participants viewed 2,400 white Western faces (1,200 female, 1,200 male, 18-35 years), each of which is described as a set of 14,319 3D shape coordinates and 2,048 $\times$ 1,072 RGB complexion pixels.

3.3.3 Experimental task

All participants read the task instructions and provided informed consent prior to testing. On each experimental trial, participants viewed a randomly generated face identity and rated it from richest to poorest on a 7-point scale (1 = richest, 4 = middle income, 7 = poorest; see Figure 3.1A right). Participants used a GUI controlled via the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007) in MATLAB R2018a to select their response from the range of response options displayed vertically to the right of the stimulus. Stimuli remained visible until response and participants were instructed to base their responses on their first impressions and to work quickly. Following response, the next stimulus appeared after an inter-stimulus-interval of 300 ms.

Each participant completed a total of 2,400 trials presented in random order across the experiment and an additional 10 practice trials before the start of the experiment. The

2,400 trials were split into four blocks of 600 trials each with a short break every 100 trials. The experiment was blocked by sex of stimulus face (2 blocks per sex of face), alternating across the experiment, with the first block randomized across participants. All face stimuli appeared on a black background in the participants' central visual field using a 1920 $\times$ 1080 resolution color-calibrated monitor. Participants viewed the stimuli at a constant distance of 72 cm maintained using a chin rest. The face stimuli (18 cm in height $\times$ 11 cm in width on average) subtended 14.25° (vertical) and 8.74° (horizontal) of visual angle. Each participant completed the experiment across three to four 1-hour-long sessions each separated by at least a one-hour break and conducted over a two- to three-day period.

3.3.4 Social class face modeling procedure

Following the experiment, I derived a model of the facial features that elicit social class perception in each participant by measuring the statistical relationship between the 3D shape and 2D complexion variations presented on each trial (represented as identity component coefficients – see Figure 3.1B, left panel) and the participant's corresponding social class ratings, using linear regression (see Figure 3.1B, center panel; see also section 6.3 in the Supplementary Material for a discussion of the appropriateness of using a linear model). Specifically, I first reverse-coded the participants' responses such that higher values denoted higher social class (i.e., 1 = poorest, 7 = richest) for ease of interpretation (higher number = higher class). Next, for each individual participant and each sex of stimulus face, I measured the statistical relationship between each of the 402 PC stimulus identity component coefficients and the participant's responses using a linear regression model (RobustFit, MATLAB 2018b). Specifically, for each PC, I regressed the stimulus identity component coefficient onto the participant's response. For example, in Figure 3.1B (center panel), the face shape variations represented by the orange color-coded identity component is positively associated with this participant's social class ratings (see orange regression line). I applied this analysis to 3D shape and 2D complexion separately, resulting in a total of 60 models [30 participants × 2 sex of stimulus face] for 3D shape and 60 models for 2D complexion for social class. Figure 3.1B (right panel) shows one example participant's resulting face models of a rich and poor female face. Note that I also re-computed each social trait face shape and complexion model in this manner as previous models (see chapter 2 – Social trait face modeling procedure) were based on ridge regression on the 3D vertex and 2D pixel stimulus information, not identity components. Models from the two analyses were highly correlated (shape: minimum r = .92, median r = 1, corrected (Bonferroni-Holm) p < .001; complexion: minimum r = .76, median r = .98, corrected (Bonferroni-Holm) p < .001). However, using the identity components here allowed for easier comparison to social class due to reduced noise.

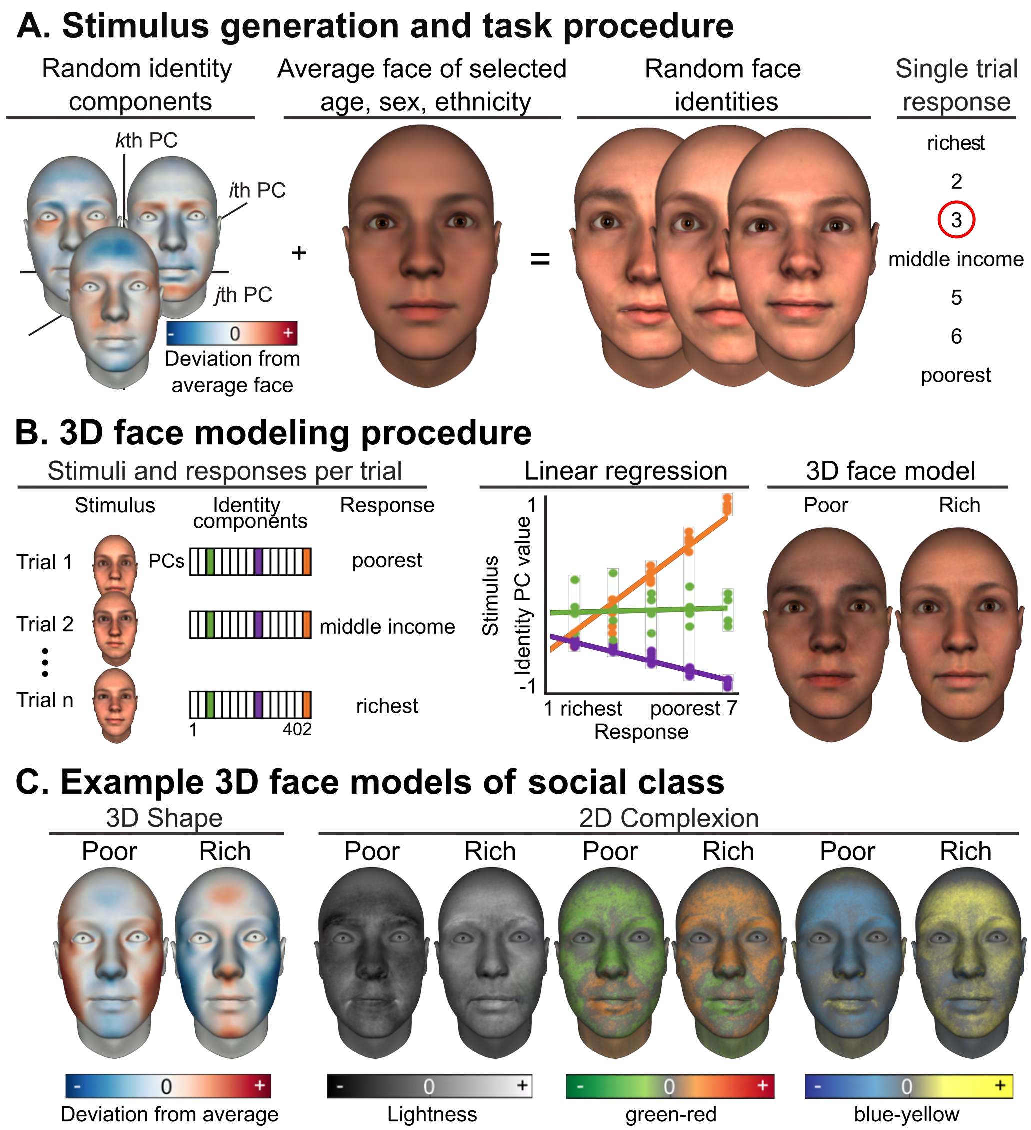

Figure 3.1. Modeling the 3D facial features that elicit perceptions of social class. (A) Stimulus generation and task procedure. A generative model of 3D face identity added random face identity components (3D shape and 2D complexion) to an average face, producing random face identities. Participants rated each face according to social class (e.g., see red circle). (B) 3D face modeling procedure. To model the facial features that drive the perception of social class in each participant, I linearly regressed the face variations (i.e., the identity PC coefficients, color-coded) on the participant's social class responses. Here, variation of the 3D face shape represented by the orange color-coded identity component is strongly associated with this participant's social class perceptions (see orange regression line). One example participant's resulting female face models are shown on the right. (C) Example face models of social class. Each face shows, for 3D shape and 2D complexion separately, the direction and magnitude of deviation from the average face, for one example participant's poor and rich female face models (see color bars at bottom). For shape, red represents positive deviation from the average, blue represents negative deviation. For complexion, colors represent the direction and magnitude of the deviation in L*a*b color space. Values are normalized across shape and complexion separately.

Figure 3.1C shows the resulting 3D shape and 2D complexion models for the same participant. Each face shows the deviation of these features from the average face (female, white, aged 18–35 years). In 3D shape, red indicates positive deviation from the average – for example, this participant's poor face model is wider than average – and blue indicates negative deviation – for example, this participant's rich face model is narrower than average. In 2D complexion, coloration indicates the deviations of lightness, green-red, blue-yellow from the average (see colorbar to bottom). For example, this participant's rich face model has lighter, redder, and yellower cheeks than average. Before further analysis of the face models, I validated each individual model of social class for 3D face shape and 2D complexion separately with methods similar to those used in chapter 2, as follows.

3.3.5 Validating 3D social class face shapes

To validate each social class 3D face shape model, a new group of participants judged the social class of faces generated from these models. Twenty participants (white, British, 10 female, 10 male; mean age = 22.80 years, *SD* = 3.49 years; mean subjective social status = 5.70, *SD* = 0.92) took part, recruited based on the same recruitment method, eligibility criteria, and pre-screening measures as described above. Using each of the 60 face shape models [30 participants $\times$ 2 sex of stimulus face], I generated seven faces ranging from poorest to richest including extrapolations, resulting in a total of 420 faces [7 generated faces $\times$ 60 models; 210 females, 210 males]. I fixed the face complexion across trials using the average complexion model. I then created all pairwise combinations of faces from each individual model (excluding identical pairs) for each sex separately, resulting in 1,260 pairs of faces [30 participant face models $\times$ 2 stimulus face sexes $\times$ 21 pairings of the 7 faces ranging from poorest to richest].

On each experimental trial, participants viewed a pair of same-sex faces generated from the same model, presented side-by-side (with the side each face appeared on randomized), and judged which of the two appeared poorer (or richer, depending on the task block) in a 2-alternative forced choice task. Each participant judged all 1,260 pairs of stimuli in each of two separate tasks – poorer or richer – conducted in blocks randomized across participants. Participants responded using a GUI by clicking on one of the faces, with the next pair of stimuli appearing after an inter-stimulus-interval of 300 ms following response. Participants were instructed to work quickly and base their judgments on their first impressions. Each participant therefore completed a total of 2,520 trials [1,260 stimulus pairs $\times$ 2 tasks; 630 female and 630 male face pairs per task]. Trials were split into four blocks of 630 trials each, with each block comprising one judgment task (poorer, richer) and one sex of face stimuli. Both whether participants completed the female or male blocks first and which judgment task they completed first were randomized. Within each block, trials appeared in random order and participants took short breaks after every 105 trials. All face stimuli were displayed in the same way as described above (see subsection 3.3.3 – Experimental task). Prior to testing, all participants provided informed consent and completed 10 practice trials. Participants completed the experiment across three to four one-hour-long sessions completed across two or three days.

To measure the validity of each of the 60 face shape models, I computed the average proportion of trials on which the participants accurately selected the richer or poorer generated face and compared this to chance performance (0.5) using a single-sample t-test and Bonferroni-Holm correction for multiple comparisons. Results showed that 56 models (93.33%, 27 females, 29 males) exceed statistical threshold ($t \geq 3.24$, corrected $p \leq .02$; see Figure 6.10A in the Supplementary Material). I retained these validated models for further analysis.

3.3.6 Validating 2D social class face complexions

I validated the social class 2D face complexion models in the same way as the social trait 2D complexion models (see chapter 2 – Validating 2D social trait face complexions). Specifically, I used a leave-one-out cross-validation approach, which tests the extent to which the complexion features of one participant can be accurately classified (i.e., as rich or poor) based on the complexion features derived for all other participants. Specifically, I iteratively trained a general linear model on 29 of the 30 participants' responses (with stimulus complexion as the predictor and participant responses as the outcome) and used the left-out participant as the test case. For each iteration, I compared the ranking consistency of the values predicted by the training model with the participant's actual ratings using Spearman's rank correlation. Following Bonferroni-Holm correction for multiple comparisons, results showed that 47 complexion models (78.33%; 22 female, 25 male) exceeded statistical significance threshold ($\rho \geq .084$, corrected $p \leq .05$; see Figure 6.10B in the Supplementary Material). I retained all validated models for further analysis.

3.3.7 Visualizing the features of the face models

To characterize and visualize the 3D shape and 2D complexion face features of the identity component based social class and social trait models, I first used the PC regression values to derive predicted identity component coefficients and converted these to vertex values (x, y, and z coordinates) for 3D shape and to pixel values in L*a*b space for 2D complexion using the generative face model used to generate the stimuli. I did this for each validated individual model (3D shape: n = 284 total [56 social class + 228 social traits]; 2D complexion: n = 267 total [47 social class + 220 social traits]).

I then proceeded as in chapter 2 – Visualizing the face features of the social trait face models. Specifically, I computed the difference between these predicted values and those representing the average face from the generative face model. For 3D shape, I subtracted the average face's x, y, and z vertex values from those of each predicted face, calculated the cosine of the angle between the model's difference from average (i.e., residual) and the vector vertical to the tangent of each vertex of the average face, and calculated the mean of the values across the validated models for each sex of stimulus face. For 2D complexion, I calculated the difference from the average face, for L, a, and b separately. Figure 3.2A shows the results (for average shape and complexion models of social traits and social class displayed side-by-side see Figure 6.22 in the Supplementary Material).

3.4 Results

A visual inspection of the face maps shows that faces judged as poor are wider, shorter, and flatter, with downturned mouth corners (see 3D shape) and darker, cooler (greener, bluer) complexions with male faces showing more ruddy (redder, yellower) noses and mouths than the average face (see 2D complexion). To illustrate how these features overlap with those of social traits, Figure 3.2B shows the features that have the same direction from the average face (mean across participants) for each social class and social trait (female faces only for illustration). Specifically, on the left, face maps show shape features of social class (exaggerated for clarity). In the middle, purple and green face maps show featural overlaps between social class and social traits for female faces. On the right, face width is shown as an illustrative example of feature comparisons between the social judgments with red showing social class features, blue hues showing social trait features, and black showing the average face. The figure illustrates that faces judged as dominant, incompetent, untrustworthy, and cold each comprise features similar to poor, including a wider and shorter face with downturned mouth corners. Similarly, faces judged as poor comprise wider than average faces similar to incompetent and cold (see cross-section on the right). In contrast, faces judged as submissive, competent, trustworthy, and warm comprise features similar to those of faces judged as rich, including narrower face width, arched eyebrows, and longer, thinner noses.

3.4.1 Mapping the relationship between the perception of social traits and social class

To formally examine these feature overlaps between social class and social traits, I proceeded in two steps. First, I replicated the finding that the perception of social class predicts (i.e., correlates with) the perception of these four social traits (Bjornsdottir & Rule,

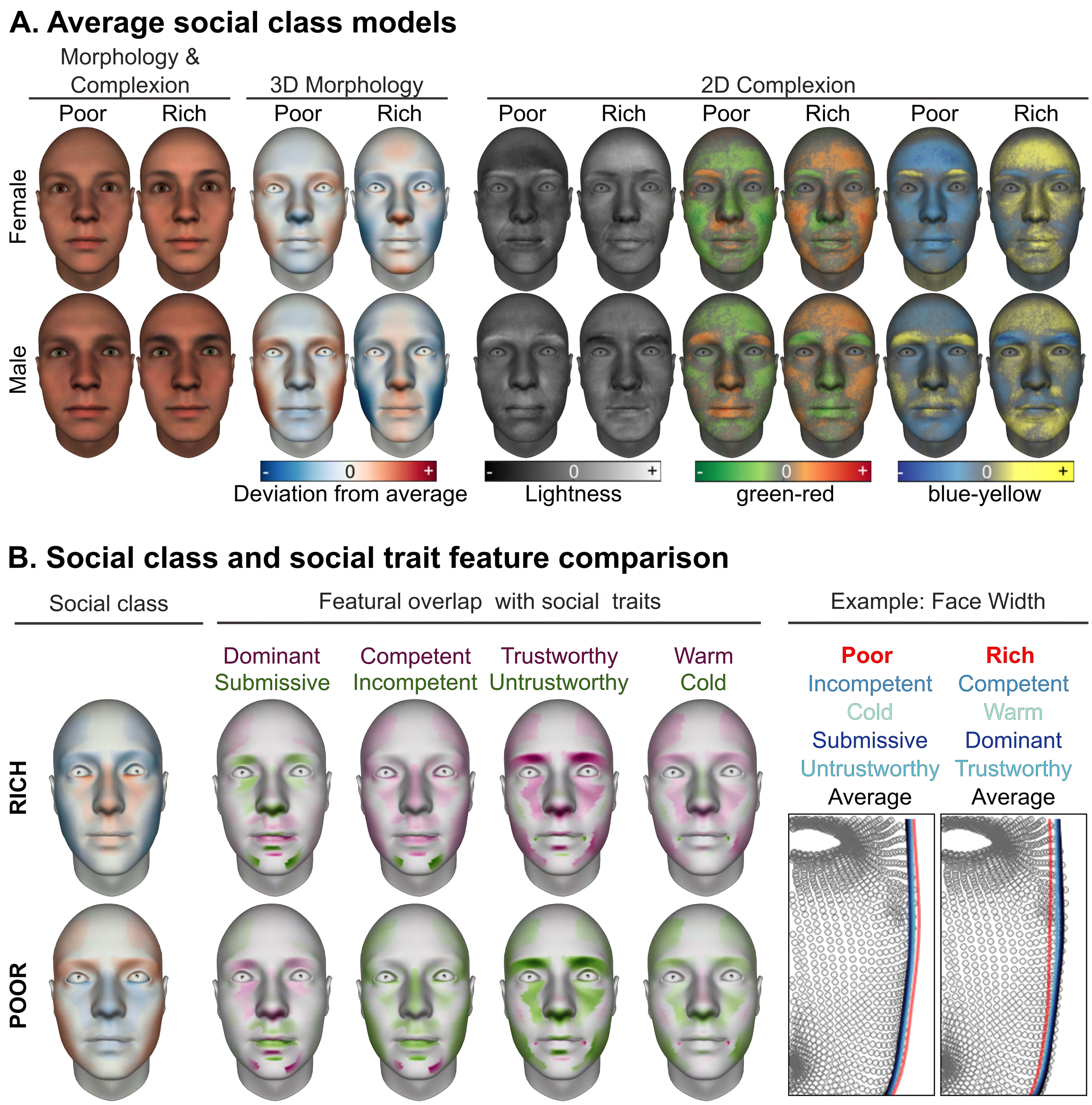

Figure 3.2. Average 3D face models of social class and overlap with social traits. (A) Faces on the left show the average predicted face models for social class and each stimulus sex. Beside it, results are shown separately for 3D shape and 2D complexion. For shape, color-coding shows the direction of deviation from the average face (red = positive; blue = negative; see colorbars on bottom; see also Figure 6.11 in the Supplementary Material for results separated into x, y, z planes). For complexion, color-coding represents the direction and magnitude of the deviation from the average face in L*a*b color space separately. Values are normalized for shape and complexion separately. See Figure 6.12 in Supplementary Material for cross-participant convergence. (B) Illustration of the featural overlap between social class and social traits. On the left, colored face maps show exaggerated social class features (female faces). Purple and green face maps show features that overlap (i.e., shape deviation from average face in the same direction) with each trait (see color coding at top). To the right, face width is shown as an example comparison between each social class (red) and stereotypically related social traits (blue hues; average face values shown in black).

2017; Bjornsdottir, 2019). To do so, I correlated the average social class rating (mean across 30 participants) of each of the 2,400 randomly generated face identities with their corresponding average rating for each social trait separately (mean across 30 participants) for each sex of stimulus face. Results showed that judgments of social class correlate positively with judgments of competence, $r(2398) = .62$, $p \leq .001$, warmth, $r(2398) = .51$, $p \leq .001$, and trustworthiness, $r(2398) = .52$, $p \leq .001$, for male and female faces, and weakly negatively correlate with judgments of dominance, $r(1198) = -.06$, $p = .03$, for male faces (see Figure 3.3). Therefore, faces rated as richer tend to be rated as more competent, trustworthy, and warm, and slightly less dominant than faces rated as poorer, mirroring existing findings (Bjornsdottir & Rule, 2017; Bjornsdottir, 2019; Diener & Biswas-Diener, 2002; Fiske et al., 2002).

This suggests that the perception of social class could be driven by a set of facial features that also drive the perception of specific social traits, thus sharing a common feature space. To test this directly, I examined the similarity of the facial features associated with the perception of social class and those associated with each social trait using two complementary analyses – Principal Components Analysis and a machine learning classifier approach. Specifically, I first applied a PCA to all validated face models of social class and social traits using the predicted identity component coefficients for the extreme ends of each social judgment – e.g., 'very competent', and 'very incompetent' – , for 3D shape and 2D complexion and for each sex of stimulus face separately. To do so, I pooled the predicted identity component coefficients (402 shape; 402 $\times$ 5 complexion) across all five social judgments (social class, four social traits), separately for each sex of stimulus face and for 3D shape and 2D complexion and applied a PCA to each. To identify the optimal number of PCs, I used the broken stick and elbow methods (see Figure 6.23A in Supplementary Material) and cross-validated the results by reconstructing the predicted faces from the PC solutions and comparing them to the original models by measuring the correlation between the 3D face information of the reconstructed faces and the 3D face information of the faces reconstructed based on all PCs (see Figure 6.23B in Supplementary Material). Together, these results yielded a four PC solution for female 3D shape models, a three PC solution for male 3D shape models, a three PC solution for female 2D complexion models, and a 2 PC solution for male 2D complexion models. I then examined the distribution of the individual participant face models within each derived PC space to evaluate the extent to which they share features and/or comprise specific face feature accents. Figure 3.4 shows the results (see also Supplementary Material – Figure 6.13C for green-red results in colorblind friendly colors).

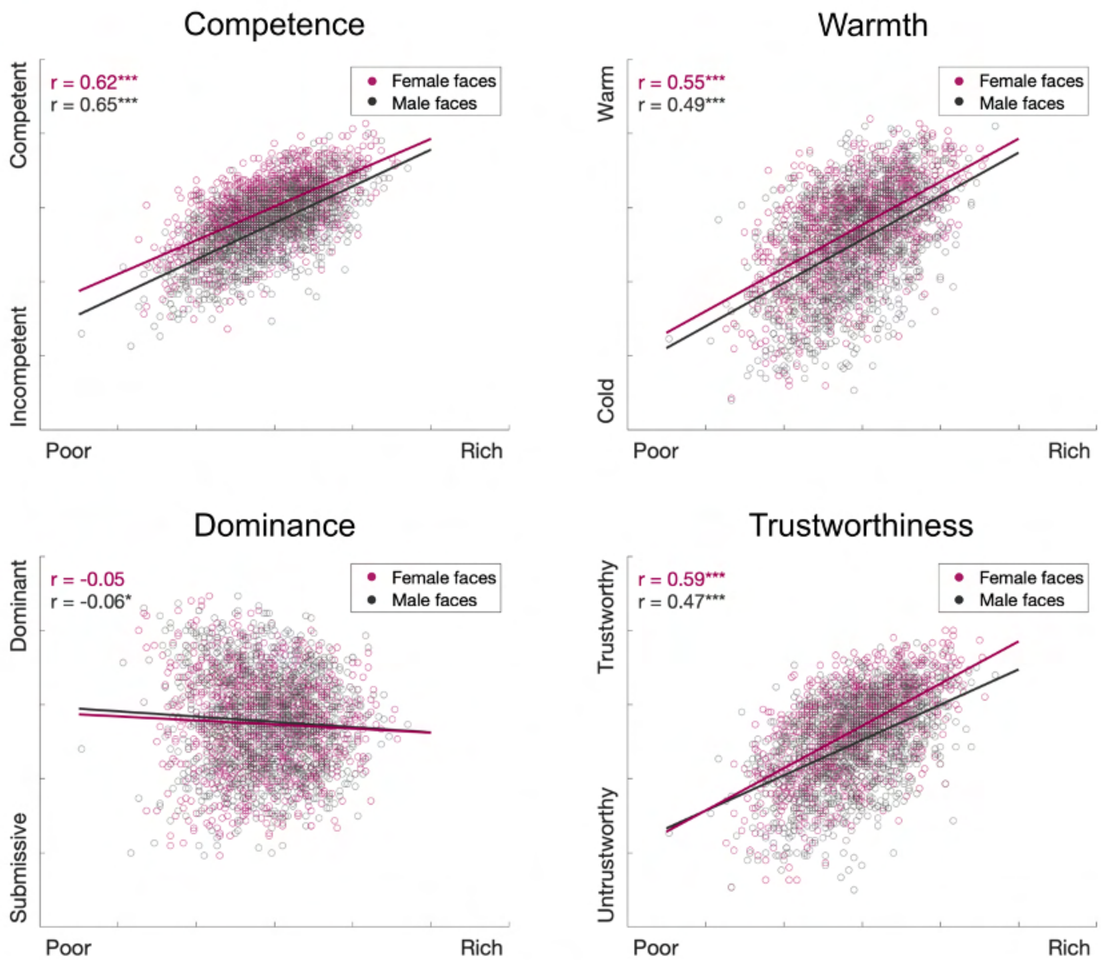

Figure 3.3. Relationship between social class and social trait face perception. Color-coded points in each plot show the average social class rating (x axis) of a randomly generated face identity plotted against its average social trait rating (y axis). Each point thus represents one of the 2,400 randomly generated face identities. Purple represents female faces, black represents male faces. Color-coded lines represent the correlations for each sex of stimulus face (correlation coefficients shown in the top left). Results show that ratings of social class correlate positively with ratings of competence, warmth, and trustworthiness with a weaker relationship with dominance. Thus, faces perceived as rich are also perceived as more competent, warm, and trustworthy, and slightly less dominant (results for female faces did not exceed threshold for statistical significance). * $p < .05$, *** $p < .001$

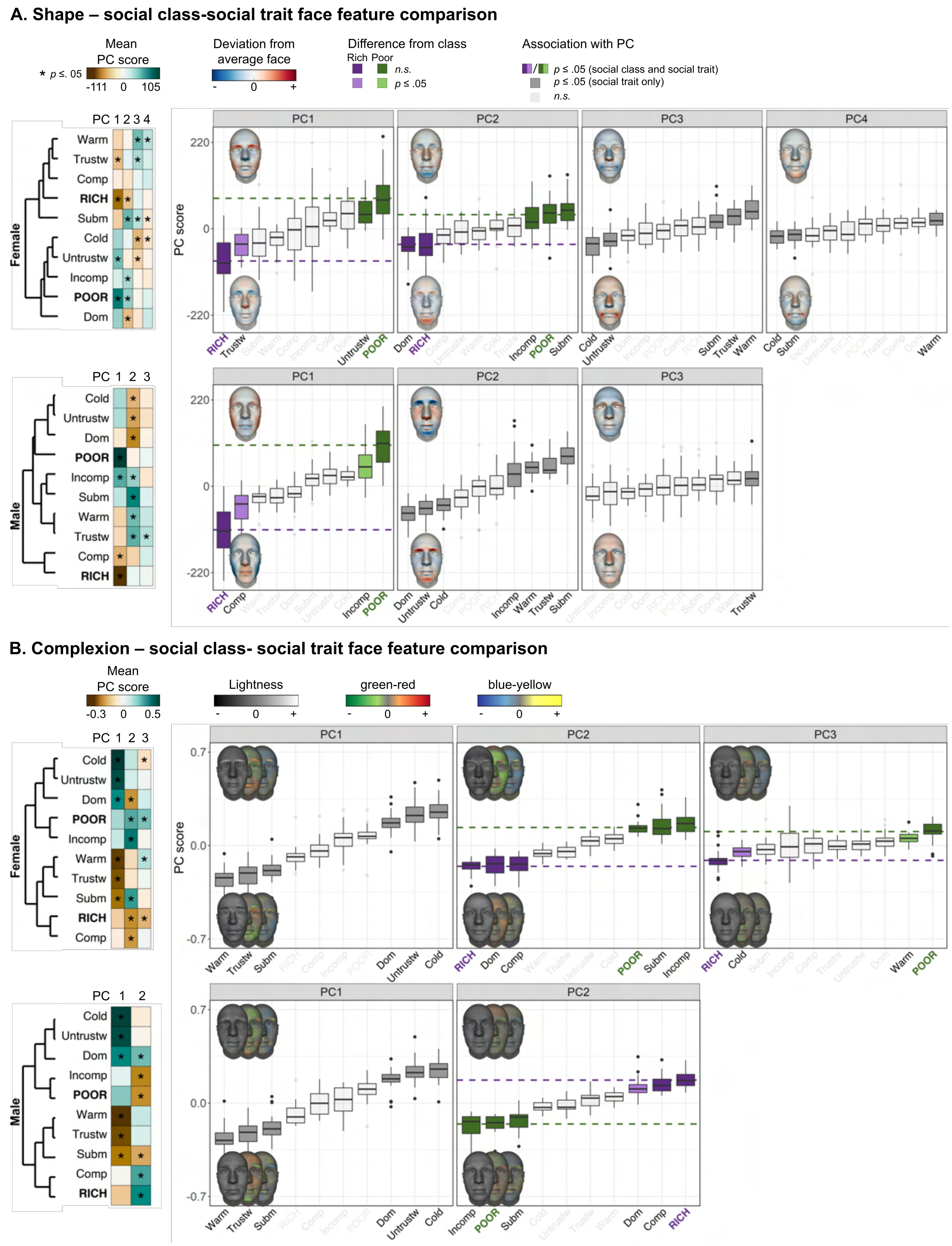
A. Shape – social class-social trait face feature comparison
Mean PC score
* p ≤ .05
-111 0 105
Deviation from average face
- 0 +
Difference from class
Rich Poor
n.s.
p ≤ .05
Association with PC
p ≤ .05 (social class and social trait)
p ≤ .05 (social trait only)
n.s.
Female
Male
PC score
PC1
PC2
PC3
PC4
B. Complexion – social class- social trait face feature comparison
Mean PC score
-0.3 0 0.5
Lightness
green-red
blue-yellow
- 0 +

Figure 3.4. Face feature comparison between the social class and social trait face models. (A) Shape – social class and social trait feature comparison. Each column represents a PC; color-coded face maps show the face features captured by each PC (red – positive deviation from average face, blue – negative deviation; see colorbar; normalized within each PC). Color-coded box plots show the distribution of PCA scores for each social judgment (see labels on x axis; points beyond 1.5 $\times$ interquartile range plotted separately; see also Figure 6.24 in Supplementary Material for PCs shown in x, y, and z planes). Purple shows comparisons between social trait face models and rich face models; green shows comparisons with poor face models (see legend at top). Darker colors indicate no statistically significant difference; brighter colors indicate a statistically significant difference (see legend at top). Dashed lines show the mean social class PC score on relevant PCs. Grayscale boxplots indicate no comparison between social class and social traits (dark gray – PC significantly associated only with social traits; light gray – not significantly associated with social traits; see legend). On left, color-coded matrices show the average PC scores for social class and each social trait, ordered according to their pairwise similarities (see dendrogram to left). Brown represents negative PC scores; blue represents positive PC scores; asterisks indicate the PC is significantly associated with the social judgment. Values normalized across sex of stimulus face and PCs, for shape and complexion separately. (B) Complexion – social class and social trait feature comparison. Results are shown using the same format as in panel A.

In each panel, color-coded face maps show the facial features represented by each PC using the same format as in Figure 3.2. Below, box plots show the distribution of per-participant PC scores for the face models of social class and each social trait (see labels on x axis), ordered via their PC scores. To evaluate the extent to which the face models of social class and the social traits comprise similar features, I compared their average PC scores using Welch two-sample*t*-tests (corrected for multiple comparisons using Bonferroni-Holm correction) and did so only if the PC is significantly associated with both social judgments, as assessed via correlations (Pearson's *r*) between the PC scores and binary social trait vectors (e.g., 'poor', 'not poor'; corrected for multiple comparisons using Bonferroni-Holm correction). Colored box plots show these comparisons (grayscale boxplots indicate no comparison). Dashed lines show the average PC score for rich and poor face models. Purple indicates comparison between social trait face models and rich face models (e.g., competent and rich), green indicates comparison with poor face models. Lighter hues indicate significant differences ($p \leq .05$ after Bonferroni-Holm correction). Darker hues indicate no statistically significant difference. To the left, color-coded matrices show the average PC scores for each social judgment, ordered according to their pairwise similarities (see dendrogram on left). Brown represents negative scores, teal represents positive scores, and asterisks indicate the PCs that are significantly associated with the face models of that social judgment.

Results reveal similarities in the facial features that drive the perception of social class and specific social traits, plus accents that distinguish social class from social trait judgments. For example, rich and poor female faces comprise face shape and complexion features (represented by PC1) that are more pronounced compared to specific social traits (see bright purple and green box plots) plus features (represented by PC2) that are more similar to specific social traits (see dark purple and green box plots). Similarly, the color-colored matrices show that, overall, the face models of social class comprise a specific combination of facial features (i.e., PCs and scores) that is distinct from all social traits. For example, rich female faces comprise a specific combination of face shape and complexion features represented by PCs 1–2 and PCs 1–3, respectively (see brown coloration and asterisks) that is not observed for any social trait.

Together, these results show that female faces judged as rich comprise face shape and complexion features that are similar to those of faces judged as dominant and competent – i.e., lowered brow, protruding chin, and a lighter and warmer complexion (see PC2 for shape and complexion) – with similar results found for complexion for rich male faces (see PC2). Equally, while rich female face models comprise features that are also observed in the face models of trustworthy – i.e., narrow faces with protruding features and upturned mouth corners (see PC1) – they are more pronounced for rich face models. Similarly, male rich and competent face models also comprise these features – narrow faces with upturned mouth corners (see PC1) – but they are more pronounced for rich models.

In contrast, female faces judged as poor comprise face shape and complexion features that are similar to those of faces judged as untrustworthy (shape PC1), incompetent, and submissive (PC2) – i.e., wider faces with flatter features, downturned mouth corners, and darker, cooler complexions. Male poor faces comprise complexion features that are similar to those of incompetent (shape PC1, complexion PC2), and submissive (complexion PC2) face models – i.e., a darker, cooler complexion. Equally, while poor female face models comprise complexion features that are also observed in the face models of warm (PC2) they are more pronounced. A similar pattern of feature pronunciation is found for poor male face models for incompetent shape features. In sum, these results show that while faces judged as rich or poor comprise face features that are similar to those found in faces judged according to specific social traits, they comprise a specific combination of facial features that distinguishes them from these social traits, including more pronounced features.

3.4.2 Classification of social class and social trait face models

To further test the distinctiveness of the social class and social trait face models, I used a machine learning approach to classify the face models by social class and the four social traits. Specifically, in a stratified 10-fold cross-validation, I trained a supervised learning classifier – Support Vector Machine (SVM) – to classify the face models reconstructed from the analyzed PCs for 3D shape and 2D complexion and for each sex of stimulus face separately. Specifically, I reconstructed 282 validated models of female shape based on 4 PCs, 286 validated models of male shape based on 3 PCs, 260 validated models of female complexion based on 3 PCs, and 274 validated models of male complexion based on 2 PCs. I converted predictor variables to z-scores within each fold to prevent data-leakage and employed automatic hyperparameter optimization for coding design, lambda, and binary learner type. In each case I trained the classifier to predict the ten social judgment labels (e.g., 'rich', 'poor', etc.) based on the reconstructed identity components (e.g., 402 identity components $\times$ n reconstructed social class and social trait models). The SVM classifier produced a mean classification error across folds of 46.7% for female shape models, 46.3% for male shape models, 29.32% for female complexion models, and 41.45% for male complexion models. Figure 3.5 shows the aggregated classifier performance across all folds. Accurate classifications of the face models would indicate that they comprise sufficiently distinct features, whereas consistent misclassifications would indicate that they do not.

In each panel, colored matrices show the face model classification performance with green indicating correct classifications – i.e., diagonal squares – and pink indicating misclassifications – off diagonal squares (see colorbar to right). Numeric values in each cell show the percentage of face models classified/misclassified. Squares outlined in orange denote misclassifications between social class and social traits above chance (i.e., 1/number of incorrect classification options; misclassifications between social traits are not highlighted). As shown by the green diagonal values, classification performance is high with some systematic misclassifications – for example, poor face models are generally misclassified as incompetent and rich faces are generally misclassified as competent (and vice versa). However, correct classifications systematically outnumber misclassifications. Together, these results show that the face models of social class comprise a specific combination of face features that is highly distinct to most social traits but shares many features with face models of competence.

In sum, by comparing the facial features of social class and social trait faces models using PCA and then testing their distinctiveness using a machine learning approach, results showed that the face features driving the perception social class comprise a specific combination of shape and complexion features that are similar but not equal to those driving the perception of social traits. The highest similarity was observed between social class

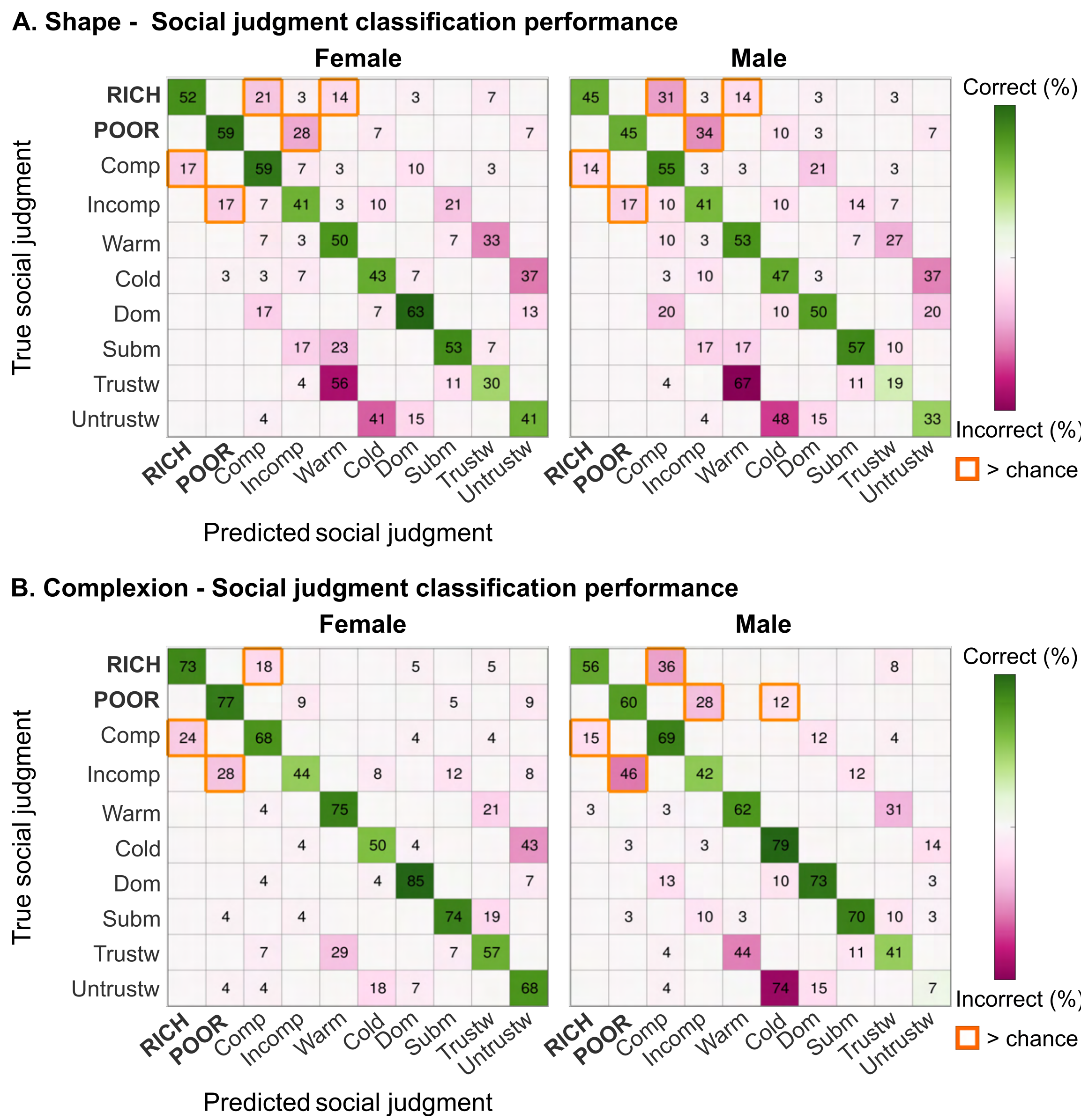

Figure 3.5. Social judgment classification performance of a supervised learning classifier. (A) Social judgment classification performance for 3D shape and (B) 2D complexion for each sex of stimulus face. Color-coded matrices show the classifications (x axis) of the face models (y axis). Green shows correct classifications – i.e., diagonal squares; pink shows incorrect classifications – i.e., off-diagonal squares (see colorbar to right). Squares outlined in orange denote above-chance misclassifications between social class and social traits. Numeric values in each cell show the percentage of face models classified/misclassified.

and competence on the basis of their face width. Thus, while many of the facial features driving the perception of social class also contribute to the perception of related social traits – here, competence, warmth, dominance, and trustworthiness – they comprise specific face features that mostly distinguish social class as a specific social perception.

3.5 Discussion

Chapter 3 addressed the central question of which facial features drive the perception of social class and whether they correspond to those that drive the perception of related social traits. In doing so, it thus provides an important explanatory element in understanding the intrinsic relationship between the perception of social class and social traits. Specifically, I used a generative model of 3D human faces, a data-driven psychophysical method, and subjective social perception to model the face features (3D shape and 2D complexion) that elicit the perception of social class. Examination and comparisons of these face models using PCA showed that the perception of social class is systematically associated with specific facial features – for example, faces judged as poor tend to be broader and shorter with more widely spaced and lower eyes, flatter and shorter noses, downturned mouths, and darker, cooler complexions. In contrast, faces judged as rich tend to be narrower with longer, more protruding features, upturned mouths, and lighter, warmer complexions.

Further, I found that these features generally correspond to those that drive the perception of social traits that are stereotypically associated with social class – for example, faces judged as poor comprise features that are also found in faces judged as submissive, incompetent, and untrustworthy and faces judged as rich comprise face features that are also found in the face models of competent, and trustworthy, and with some variation according to sex of stimulus face. These results correspond with existing results showing that judgments of social class correlate with judgments of social traits from faces (Bjornsdottir & Rule, 2017; Bjornsdottir, 2019), as replicated here (see Figure 3.3). The analysis also revealed systematic differences in the facial features associated with the perception of social class and social traits, whereby the social class face models comprise a specific combination of facial features that is either distinct from or more pronounced than those of the social trait face models. To test this directly, I used a supervised learning classifier to classify the face models according to social class and the four social traits – results showed high classification performance but with systematic confusions between social class and competence. Together, these results show that the facial features that drive the perception of social class are related to those of social trait, and in particular competence perception but nevertheless comprise a set of features that is somewhat distinct from that of each social trait.

These results highlight the close relationship between the perception of social class and specific social traits, which reflects well-documented social class stereotypes (e.g., Bjornsdottir & Rule, 2017; Cuddy et al., 2008; Fiske et al., 2007). Specifically, I found that faces judged as rich comprise features that are also found in faces judged as competent, and trustworthy and those judged as poor comprise features that are also found in faces judged as incompetent, and untrustworthy. Dominance, however, showed a conflicting pattern of overlapping features, such that some features eliciting perceptions of dominance also elicit perceptions of high social class (PC1) and others of low social class (PC2), helping to explain the near-zero correlation between judgments of social class and of dominance. Equally, perceived dominance does not always lead to perceived high status (Cheng et al., 2013), and among women, greater femininity (which appears more submissive) may be associated with higher social class (Bjornsdottir, 2019). These findings expand upon existing work showing that conceptually similar person judgments correlate with overall facial feature similarity (e.g., Bin Meshar et al., 2021; Stolier et al., 2018, 2020) by demonstrating (1) which conceptually related social judgments overlap perceptually, (2) their degree of overlap, but also (3) uniquely pinpointing the sources of these overlaps in the face.

However, although the perception of social class and of specific social traits are each driven by similar facial features, the specific combination of these facial features can be distinguished, including through variations in feature pronunciation. These results suggest that while certain facial features can elicit related social perceptions, the specific combination of facial features driving the perception of social class is mostly unique to this particular social attribute. These results therefore suggest that the perception of social class is to some degree distinct from the perception of social traits and cannot be reduced to any specific social trait although it shares particularly many features with competence perception.

These results also contribute to the wider literature on face perception, including overgeneralization effects (Zebrowitz, 2017; Zebrowitz & Montepare, 2008). Specifically, many of the features driving social class perception are also tied to the perception of other social attributes such as fitness, emotion, babyfacedness, and familiarity. For example, poor faces comprise broader features and cooler complexions which might indicate higher face adiposity and are associated with poorer health (Henderson et al., 2016; A. L. Jones, 2018; Little, 2014; Said & Todorov, 2011). Similarly, downturned mouths are associated with negative emotions (e.g., Jack et al., 2016) and shorter chins and lower eyes correspond with infantile features (Berry & McArthur, 1986), which in turn are associated with perceptions of lower competence (Berry & McArthur, 1986). Finally, darker complexions could reflect ethnic outgroups for white perceivers (Blair et al., 2004).

In contrast, face features driving perceptions of rich social class standing such as narrower face width and lighter and warmer toned complexions also relate to higher health and attractiveness perceptions (Henderson et al., 2016; A. L. Jones, 2018; Little, 2014; Said

& Todorov, 2011). Crucially, attractiveness and health perceptions relate both to social class (Bjornsdottir & Rule, 2017; Bjornsdottir, 2019) and social trait judgments (Jaeger et al., 2018; Tsankova & Kappas, 2016). Similarly, upturned mouth corners relate to positive emotions (e.g., Jack et al., 2016) and more positive-looking faces in turn appear friendlier (Wolffhechel et al., 2014) and more extraverted (Borkenau et al., 2009), agentic, and communal (Walker & Vetter, 2016). Finally, nose length is associated with youthfulness/attractiveness (Vernon et al., 2014). Such associations are reflected both in general social class stereotypes such as competence and well-being (Durante et al., 2017; Varnum & Denson, 2013) and ground truth social class correlates including mental and physical health (e.g., Adler et al., 1994; Marmot et al., 1991) that laypeople often internalize (e.g., Diener & Biswas-Diener, 2002).

In sum, chapter 3 provided new insights into human social face perception by showing that judgments of social class are related to but distinguishable from those of social traits and are driven by a specific combination of facial features.

Chapter 4

Social trait facial expressions form two clusters that may reflect approach and avoidance signals

4.1 Chapter abstract

Emotion overgeneralization theory posit that resemblance of neutral faces to emotional facial expressions drives social trait inferences. Similarly, humans make stable social traits inferences from dynamic facial expressions. However, the facial movements that drive such trait inferences and how they relate to emotional facial expressions and neutral social trait face shapes are not well understood. In this chapter, I used data-driven psychophysical techniques to derive dynamic facial expression models of four key social trait dimensions – dominance, competence, trustworthiness, and warmth – and compared them to emotional facial expressions. Clustering of these models revealed that social trait facial expressions form two clusters that are related to happy and disgusted facial expressions, respectively. However, emotional features did not fully account for social trait features in either the facial expressions or the neutral face shapes. These results support theories about the importance of approach and avoidance in social trait perception but call into question the emotion overgeneralization hypothesis.

4.2 Introduction

As shown by chapter 3, social traits perceptually overlap with other key social judgments, such as social class. Notably, many of the shared facial features between social traits and social class are also implicated in other social perceptions such as emotion perception (Jack et al., 2016; Zebrowitz & Montepare, 2008). These data therefore lend support to the emotion overgeneralization effect – an inference of social attributes from facial features resembling emotion (Zebrowitz, 2017). Machine learning based approaches successfully demonstrated that correlations between social trait and emotion inferences from neutral faces stem from such face feature similarities (Albohn & Adams, 2021; Said et al., 2009). In particular, positively valenced social trait perceptions, such as trustworthiness are, at least in part, driven by facial features resembling positively valenced emotions such as happiness (Albohn & Adams, 2021; Said et al., 2009). In contrast, threat-related social traits, such as dominance, are associated with facial features of negatively valenced emotions, such as anger (Albohn & Adams, 2021; Said et al., 2009). Indeed, emotional facial resemblance is one of the strongest predictors of social traits judgments from neutral faces (Jaeger & Jones, 2022). Yet it is unclear which specific facial features drive these overgeneralization effects. For instance, while it is likely that smile resembling features drive similarity to happiness (Gill et al., 2014), shared features between dominance and anger are less easily identifiable.

Moreover, people not only infer social traits from neutral faces but also from facial expressions (Gill et al., 2014; Knutson, 1996), in particular from emotional facial expressions (Knutson, 1996). For example, angry facial expressions are judged as dominant while happy facial expressions are judged as trustworthy (Hareli et al., 2009; Knutson, 1996). However, there is little work showing the specific facial movements that drive social trait perceptions or how these relate to emotional facial expressions. Nevertheless, Gill et al. (2014) found that dynamic facial expressions driving the perception of trustworthiness and dominance comprise facial movements from several different emotions. This suggests that perceptions of social traits are not associated with any one emotion but instead are more closely associated with general affective messages. Similarly, it remains an open question how social trait facial expressions relate to neutral social trait face shapes. If both are overgeneralizations from emotional expressions, social trait facial expressions should be a mere exaggerations of the neutral face shapes and be closely associated with emotional expressions.

This chapter investigates this mapping between facial expressions that elicit social trait judgments and (1) emotional facial expressions and (2) social trait face shapes, using methods similar to those in chapter 2 and chapter 3. Specifically, using reverse correlation, I model the specific facial movement patterns associated with the perception of the four key social trait dimensions – dominance, competence, trustworthiness, and warmth. I then compare these social trait facial expression models to (1) emotional facial expressions and

(2) the social trait face shapes derived in chapter 2 before (3) comparing social trait face shapes with emotional expressions.

4.3 Data-driven modeling of facial expressions that elicit perceptions of social traits

4.3.1 Participants

Twenty-three white Western participants (14 female, 9 male; mean age = 24.0 years, *SD* = 3.15 years) completed the facial expression rating task. Participants had minimal exposure to non-Western cultures as assessed via questionnaire (see section 6.1 – Participant questionnaire in the Supplementary Material for details) to reduce the impact of cross-cultural differences on perception. Participants had normal or corrected to normal vision and reported no symptoms of synesthesia, or psychological, psychiatric, or neurological conditions that can affect face-processing, as per self-report. Participants gave written informed consent and received £6/hour for participation. The experiment was approved by the University of Glasgow College of Science and Engineering Ethics Committee.

4.3.2 Stimuli

I produced random facial expression stimuli using a dynamic facial expression generation platform (GFG; Yu et al., 2012) as shown in Figure 4.1. For each stimulus, the GFG randomly selected between one and five Facial Action Units (AUs) from a set of 42 AUs, each representing an individual facial muscle movement (Ekman & Friesen, 1978, Figure 4.1 left). For each AU separately, the GFG then assigned random but biologically plausible values for each of 7 temporal parameters (onset latency, acceleration, peak latency, peak amplitude, sustainement, deceleration, and offset latency). Using this method, I created a set of 2,000 different, random, dynamic facial expressions (see section 6.6 in the Supplementary Material for an exploration of how many trials are needed to derive stable facial expression models). I presented these facial expressions on a set of 2,000 photo-realistic 3D face identities (1,000 female, 1,000 male; 18-35 years, white Western) randomly generated using the same Generative Face Grammar (Zhan et al., 2019) and procedure as in chapter 2 and chapter 3.

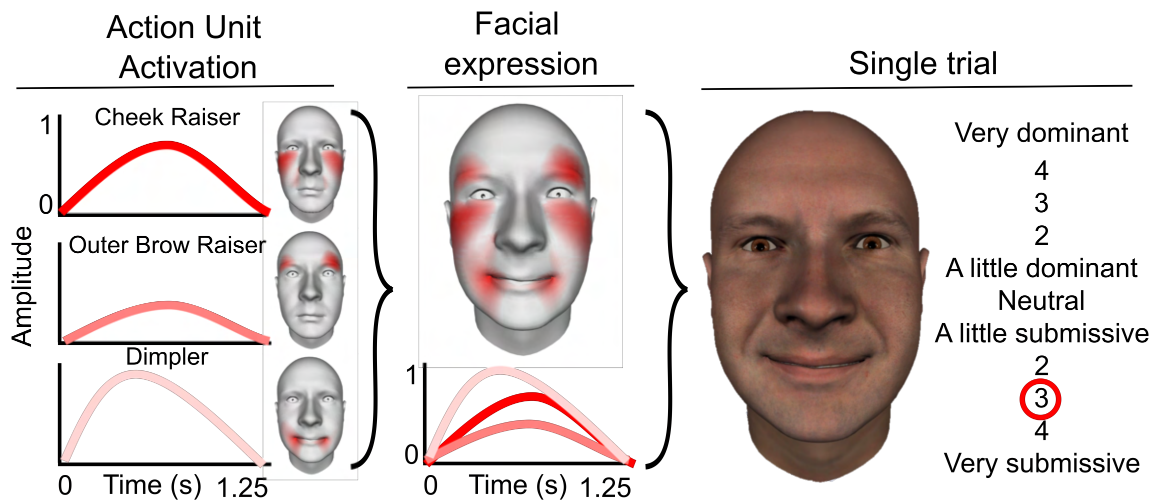

Figure 4.1. Modeling facial expressions that elicit social trait perceptions. On each trial, a generative face grammar selected random Action Units and assigned each one random temporal parameters illustrated as three example AUs' activation curves on left. Colored face maps show the corresponding facial movements. These AUs combine to a single dynamic facial expression (middle panel) of length 1.25 s. Participants rated these facial expressions on each one of four social trait dimensions (right).

4.3.3 Experimental task

To obtain a set of facial expressions associated with the perception of key social traits, each participant rated each facial animation on each of the four social trait dimensions - dominance, competence, trustworthiness, and warmth (for an example see Figure 4.1 right). Specifically, on each trial, participants viewed a random facial animation and rated it on a given social trait on a scale from 1 (e.g., 'very submissive') to 11 (e.g., 'very dominant') with 6 being the mid-point labeled 'neutral'. The expansion of the scale from 7 points in chapter 2 to 11 points ensured higher methodological similarity to the procedure used to obtain emotional facial expression models (for details see Jack et al., 2014, but also section 6.7 in the Supplementary Material).

Each participant rated each facial animation (n = 2,000) on all four social traits in separate, randomized blocks. Each block consisted of 200 trials during which the sex of stimulus face and target social trait remained constant. In total, each participant therefore completed 8,000 trials across 40 separate blocks during between nine and eleven separate 1-hour sessions [1,000 trials $\times$ 2 sex of stimulus face $\times$ 4 social trait dimensions]. To minimize fatigue, participants took breaks between blocks and sessions and completed no more than three sessions within a day. Participants responded using a GUI on a 17-inch flat panel monitor. I presented each animation in the central visual field for 1.25 s at a constant viewing distance of 70 cm, subtending 14.36° (vertical) and 8.75° (horizontal) of visual angle. I controlled the experiment using MATLAB R2018a and the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007).

4.3.4 Facial expression modeling procedure

To model the dynamic facial expressions associated with each social trait, I extracted the AUs significantly associated with the perception of each social trait. Specifically, for each sex of stimulus face, social trait (dominant, competent, trustworthy, warm, submissive, incompetent, untrustworthy, cold), participant, and AU separately, I first binarized the responses (e.g., 'submissive', 'not submissive') and counted how frequently each AU appeared for each response. Next, I performed a Monte Carlo simulation (one-tailed) with 1,000 iterations, shuffling responses at each iteration and re-counting AU frequencies. I retained those AUs with frequencies above 95% of the distribution of random frequencies thereby obtaining a 42-dimensional vector of on/off values for each AU, social trait, participant, and sex of stimulus face.

Next, to model the corresponding dynamic movements, I performed an constrained least squares regression between the response vector (i.e., social trait ratings) and each of the seven temporal parameters, for each sex of stimulus face, participant, social trait, and significant AU separately. I thus derived a set of 184 social trait facial expression models [23 participants $\times$ 4 social traits $\times$ 2 sex of stimulus] each of size 42 (AU) $\times$ 8 (on/off + 7 temporal parameters). Finally, to derive aggregate facial expression models (mean across participants), I first determined which AUs passed the population prevalence threshold (Ince et al., 2021) for each social trait and sex of stimulus face separately. For AUs which passed population prevalence, I then took the mean AU temporal parameters across participants for whom the AU was significant. Figure 4.2A shows the resulting facial expression models for each sex of stimulus face and each social trait. Specifically, colored face maps show the AUs that passed population prevalence (mean amplitude across participants) for each social trait in red. Below, colored matrices show participant counts per AU with more saturated red corresponding to higher numbers of participants (see colorbar to right).

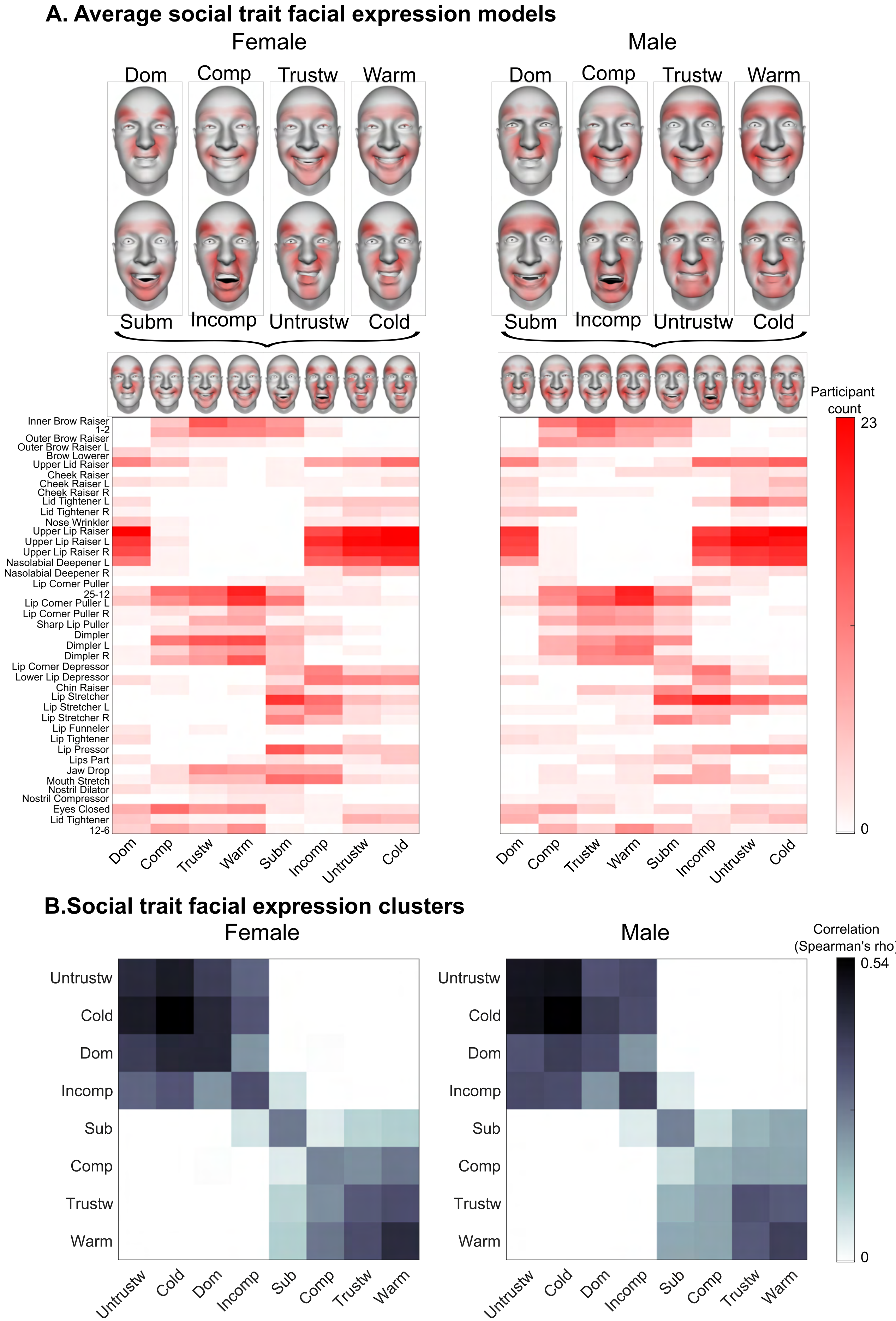
A. Average social trait facial expression models
Female
Male
Dom
Comp
Trustw
Warm
Subm
Incomp
Untrustw
Cold
Participant count
23
0
Inner Brow Raiser 1-2
Outer Brow Raiser
Outer Brow Raiser L
Brow Lowerer
Upper Lid Raiser
Cheek Raiser
Cheek Raiser L
Cheek Raiser R
Lid Tightener L
Lid Tightener R
Nose Wrinkler
Upper Lip Raiser
Upper Lip Raiser L
Upper Lip Raiser R
Nasolabial Deepener L
Nasolabial Deepener R
Lip Corner Puller 25-12
Lip Corner Puller L
Lip Corner Puller R
Sharp Lip Puller
Dimpler
Dimpler L
Dimpler R
Lip Corner Depressor
Lower Lip Depressor
Chin Raiser
Lip Stretcher
Lip Stretcher L
Lip Stretcher R
Lip Funneler
Lip Tightener
Lip Pressor
Lips Part
Jaw Drop
Mouth Stretch
Nostril Dilator
Nostril Compressor
Eyes Closed
Lid Tightener 12-6
B.Social trait facial expression clusters
Female
Male
Correlation (Spearman's rho)
0.54
0
Untrustw
Cold
Dom
Incomp
Sub
Comp
Trustw
Warm

Figure 4.2. Average social trait facial expression models and clusters. (A) Average social trait facial expression models. Colored face maps show significant AUs (assessed via population prevalence) per social trait and sex of stimulus face. Below colored matrices indicate the number of individual participant models that included each AU where more saturated red shows higher participant numbers (see colorbar to right). (B) Social trait facial expression clusters. For female and male faces separately, colored matrices show the cross-correlation (Pearson's r; mean across participants) between social trait facial expression models (on/off binary AU vectors). Darker colors show higher correlations, lighter colors lower correlations (see colorbar to right). Both matrices are ordered according to female face hierarchical clustering of the pair-wise mean correlation values.

4.3.5 Validating social trait facial expression models

To validate the social trait facial expression models I used a Leave-One-Out-Cross-validation. Specifically, for each sex of stimulus face and social trait, I trained 23 general linear models on the stimulus data [1,000 trials $\times$ 42 AUs (on/off)], each one to predict the responses of 22 out of the 23 participants. I then tested each model against the remaining response data of the left-out participant. Finally, I correlated (Pearson's r), for each sex of stimulus face, social trait, and participant separately, the predicted with the actual responses. I defined models with correlations of $p \leq .05$ after Bonferroni-Holm correction to be validated. Using this procedure, the majority of models (n = 178, 97.8%; 89 female, 89 male; 46 dominance, 41 competence, 45 trustworthiness, 46 warmth) validated and were retained for further analysis (see Figure 6.25 in the Supplementary Material for individual model validation performances).

4.4 Results

Visual inspection of Figure 4.2A suggests that (1) social traits may cluster into fewer sub-clusters based on their facial expression features compared to their face shape features and (2) social trait facial expressions differ only little between female and male faces. Specifically, submissive, competent, trustworthy, and warm comprise AUs such as Lip Corner Puller and Lip Stretcher, and, for female faces, Eyes Closed. In contrast, dominant, incompetent, untrustworthy, and cold comprise AUs such as Upper Lip Raiser and Nasolabial Deepener. To test these observations formally, I first cross-correlated (Pearson's r) all binary social trait facial expression models (42 AU on/off; n = 184 models [23 participants $\times$ 8 social traits]). Next, I clustered the resulting average correlation values (mean across participant-wise comparisons) based on hierarchical clustering. Figure 4.2B shows the resulting correlation matrices, ordered via the female face hierarchical clustering results. Darker colors denote higher correlations, lighter colors denote lower correlations. For example, untrustworthy, cold, and dominant cluster together tightly based on their facial expression features with incompetence also forming part of this cluster but with lower correlations. Similarly, submissive, competent, trustworthy, and warm cluster together, although trustworthy and warm correlate more strongly with each other than with either submissive or competent. These patterns appeared highly similar between female and male faces and, correspondingly, their cross-correlation values (participant-wise comparisons) correlated highly (r = .78, p < .001).

4.4.1 Social trait facial expression clusters

To test whether these two main clusters of social trait facial expression models corresponded to particular emotional facial expressions, I used a supervised machine learning approach. Specifically, for each sex of stimulus face, I trained a supervised learning classifier (SVM) to classify emotional facial expressions (happy, surprised, fearful, disgusted, angry, sad) based on their 42-dimensional binary vector of on/off AU values (360 models per sex of face [6 emotions $\times$ 60 participants]; see Jack et al. (2014) and section 6.7 in the Supplementary Material for how emotion models were derived). I then tested the classifier on the the social trait facial expression models (184 models [8 social traits $\times$ 23 participants]), thereby gaining a predicted emotion label for each social trait facial expression model. Finally, I clustered the social traits and emotions based on these classification values (pairwise Euclidean Distance) using hierarchical clustering. Figure 4.3A shows the resulting classification matrices, ordered via hierarchical clustering that is displayed as dendrogram trees.

Results revealed that, for both female and male faces, the two social trait clusters were most frequently classified as happy and disgusted. Specifically, submissive, competent, trustworthy, and warm facial expression models (cluster 1) were classified as happy and, to a lesser degree, as sad. In contrast, dominant, incompetent, untrustworthy, and cold facial expression models (cluster 2) were classified as disgusted and, to a lesser degree, angry. Finally, submissive facial expression models were frequently classified as surprised and fearful. I further explored which Action Units drive these classifications with an ablation analysis. Specifically, I repeated the machine learning approach described above 42 times, removing a different AU at each iteration and obtaining classification results based on these reduced sets of AUs. I then subtracted from these new classification results (i.e., count of classification as a given emotion) the classification results based on the full set of Action Units. This yielded a measure of which AUs drove classifications of each social trait as a given emotion. Negative values indicated that removal of the AU decreased classifications as a given emotion, suggesting it had positively contributed to classifications when included. The results are shown in Figure 4.3B.

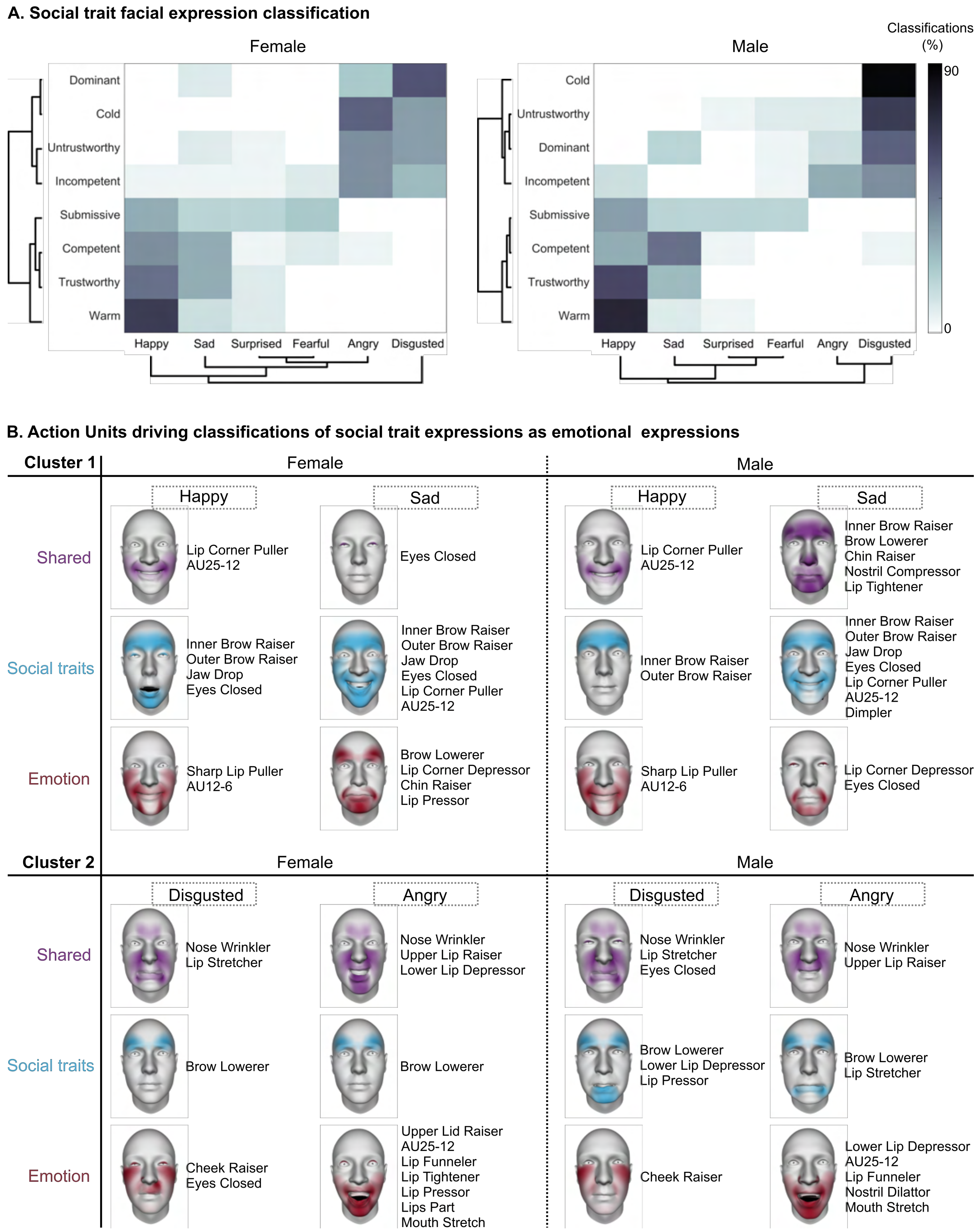

Figure 4.3. Social trait facial expression clusters and associated Action Units. (A) Social trait facial expression classification. Colored matrices show the percentage of social trait facial expression models classified as each emotion for each sex of stimulus face separately. Darker colors indicate higher percentages (see colorbar to right). Axes are ordered via hierarchical clustering of the pairwise distances (Euclidean) between these classification percentages (see dendrograms). (B) Action Units driving classifications of social trait expressions as emotional facial expressions. Purple face maps in the top row show AUs that drive classifications of social traits in each cluster as a given emotion (see also AU labels to right). Below, cyan face maps show AUs associated with social traits in this cluster but which do not drive classification results. Red face maps similarly show AUs associated with the emotion but which do not drive classification results.

Purple face maps in Figure 4.3B show AUs that drive classifications of social traits in each cluster as a given emotion. For example, social traits in cluster 1 (submissive, competent, trustworthy, warm) are classified as happy on the basis of Lip Corner Puller and AU25-12 (Lips Part + Lip Corner Puller). Below, cyan face maps indicate AUs that are significantly associated with the cluster's social traits (> population prevalence for at least two of the four social traits) but not the emotion. Red face maps show AUs significantly associated with the emotion (> population prevalence) but not the social traits. Note that in some cases, AUs were significant for both social traits and the respective emotion but did not drive classification results (see Figure 6.26 in the Supplementary Material for full results for each social trait).

Results revealed that social trait facial expression models in cluster 1 – submissive, competent, trustworthy, and warm – were classified as happy on the basis of Lip Corner Puller and AU25-12 (Lips Part + Lip Corner Puller). However, happy facial expressions further comprised Sharp Lip Puller and AU12-6 (Lip Corner Puller + Cheek Raiser) while the social traits additionally comprised Inner and Outer Brow Raiser (female and male faces), Jaw Drop, and Eyes Closed (female faces only). Additionally, social trait models in cluster 1 were classified as sad based on Eyes Closed for female faces but a wider set of AUs for male faces, including Brow Lowerer, Chin Raiser, and Nostril Compressor. In contrast, social trait facial expression models in cluster 2 – dominant, incompetent, untrustworthy, and cold – were classified as disgusted and angry in particular on the basis of Nose Wrinkler (disgusted and angry), Lip Stretcher (disgusted), and Upper Lip Raiser (angry). However, the social trait models in this cluster additionally comprised the AUs Brow Lowerer (female and male), Lower Lip Depressor, Lip Pressor, and Lip Stretcher (male). In contrast, disgusted was associated with Cheek Raiser and angry with a wider range of AUs including Upper Lid Raiser and AUs involved in the opening of the mouth such as Mouth Stretch.

4.4.2 Social trait shape features resembling facial expressions

The above results show that social trait facial expressions form two clusters that are related to happy and disgusted facial expressions, respectively. This stands in contrast to the four clusters chapter 2 identified for social trait face shapes, suggesting that there is no one-to-one mapping from social trait shape features to social trait facial expression features. To test this formally, and to identify the specific facial features that drive facial resemblances between neutral social trait face shapes and different facial expressions, I derived a measure of which specific individual Action Units correlated with each social trait face shape as illustrated in Figure 4.4.

Specifically, for each individual AU, I obtained a 3D face shape representation of the facial movement (at constant amplitude) using the GFG and obtained their deviation to

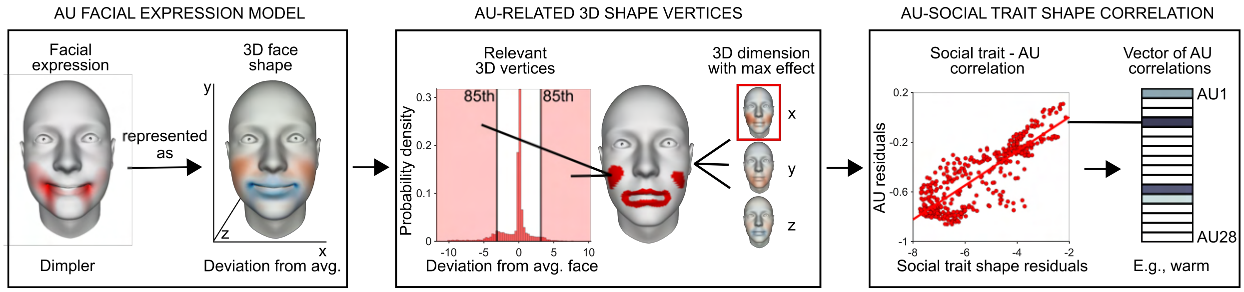

Figure 4.4. Deriving AU-representations of social trait shapes. Each panel represents a sequential step for deriving AU-representations of social trait face shape models. Left – individual AUs (n = 28 bilateral) are represented as 3D face shapes (extracted using the GFG), allowing for the calculation of the deviation to the average face. Middle – For each AU, I identified the most relevant 3D shape vertices, i.e., vertices with deviation values above the 85th or 95th percentile (depending on the magnitude of the AU's movement, e.g., Lid Raiser versus Lip Corner Puller) of the 14,319 vertices. Out of these vertices, I identified the dimension (x, y, z) with the highest absolute effect size (mean deviation from the average face). Right – For each sex of stimulus face, AU, and social trait, I correlated these shape values (residuals from the average face) and thus obtained a 28 (bilateral AUs) $\times$ 1 vector of correlation values for each individual social trait.

the average face values (Figure 4.4 left; see also subsection 2.3.7 for details on deviation values). Next, for each AU, I identified vertices above the 85th or 95th percentile (depending on the magnitude of the AU's movement, e.g., Lid Raiser versus Lip Corner Puller) across all 3D shape vertices' (n = 14,319) deviation values as most relevant to that AU (Figure 4.4 middle; see Figure 6.27 for resulting vertex models for each AU). For this, I focused on bilateral AUs (n = 28) by collapsing across lateralized AUs (e.g., Brow Lowerer Left, Brow Lowerer Right) and I retained these AUs for all further analyses to avoid, where possible, comparing the same features multiple times.

Having identified relevant 3D shape vertices for each AU, I then correlated the 3D shape vertex values of these relevant vertices between each AU and each social trait face shape, for each participant and sex of stimulus face separately (Figure 4.4 right). I corrected for multiple comparisons using Bonferroni-Holm correction and identified population level AUs via population prevalence (Ince et al., 2021). This yielded a vector of 28 correlation values for each social trait shape and sex of stimulus face, detailing the Action Unit pattern the social trait face shape features highly correlated with above population threshold. Figure 4.5 shows the resulting correlated facial features.

Red areas on each face map in Figure 4.5Ai and Bi show the face features of each AU that significantly correlated with that social trait for female (Ai) and male (Bi) faces separately. Below each face, I list the six AUs that most strongly correlated with face features of

Figure 4.5. Social trait shapes represented as Action Unit patterns. (Ai) Social trait face shape Action Unit patterns - Female. For each social trait (female faces), face maps show the shape face features (red) corresponding to AUs that significantly correlated with that social trait's face shape, derived as illustrated in Figure 4.4. Below, significantly correlated AUs are indicated, ordered by strength of correlation (top – highest correlation; highest six correlations only; see also Figure 6.28 for all correlation values). Arrows and labels indicate specific AUs that were also significantly associated with different emotions (assessed via population prevalence). For example, female dominant face shape models correlated with Lip Corner Depressor which is also common in sad facial expression models. Colored frames around each face indicate the correlation between the full social trait shape AU pattern (28 $\times$ 1 AU correlations) and the corresponding social trait facial expression AU pattern (proportion of participants per AU). Green corresponds to positive, pink to negative, and black to zero correlations. (Aii) Trait shape-emotion correlation - Female. Correlation between the social trait shape AU pattern as in Ai and the AU patterns of each of the six basic emotions. Green corresponds to positive correlations, pink to negative, and white to zero correlations. (Bi) Social trait face shape Action Unit patterns - Male. Results for male faces are shown as in Ai. (Bii) Trait shape-emotion correlation - Male. Results for male faces are shown as in Aii.

each social trait (top row = highest correlation). For example, female dominant face shapes comprised shape features similar to Lip Corner Depressor, Brow Lowerer, Upper Lid Raiser, and Nose Wrinkler. Arrows and labels around each face indicate which emotions each of these AUs was significantly associated with (assessed via population prevalence threshold). For example, Nose Wrinkler is significantly associated with angry and disgusted. Additionally, colored frames around each face map indicate the correlation (Pearson's *r*) between AU activation patterns of the social trait face shape and corresponding social trait facial expression. Green indicates positive correlations, pink negative correlations and black zero correlation. For instance, female dominant face shape models positively correlated with female dominant facial expression models. Finally, on the right, Figure 4.5Aii and Bii show correlations between AU patterns of each social trait face shape and each of the six basic emotions. Here, green indicates positive correlations, pink negative correlations, and white zero correlations.

Figure 4.5 reveals that each social trait's face shape is systematically associated with specific Action Units. In particular, submissive, trustworthy, and warm face shape models all correlated with six or more AUs that were mostly associated with smiles (e.g., Lip Corner Puller, AU12-6, Mouth Stretcher) but also with upward eyebrow movement (Inner and Outer Brow Raiser) and lengthening of the chin (Jaw Drop). In contrast, dominant, untrustworthy, and cold face shape models correlated with the opposite eyebrow (Brow Lowerer) and mouth movement (Lip Corner Depressort), as well as Upper Lid Raiser, an AU that is also associated with angry but also surprised and fearful facial expressions. Each of these social traits' AU pattern also correlated positively with its corresponding facial expression AU pattern. In particular trustworthy, warm, and dominant (male faces) showed high similarity between their shape and facial expression models, suggesting the facial expression models comprised features that were exaggerations of the neutral face shape models. However, competent and incompetent stood out through their negative correlation between shape and facial expression models (with the exception of male competence). Competent face shape models correlated with smile-related AUs such as Lip Corner Puller but also with Nostril Compressor and Jaw Drop for female faces and Brow Lowerer and Nose Wrinkler for male faces. In contrast, incompetent face shape models correlated with AUs corresponding to a downward movement of the mouth (Lip Corner Depressor, Lower Lip Depressor) but also a raising of the eyebrows.

Furthermore, these social trait shape AU patterns also correlated systematically with different emotional facial expressions (Figure 4.5Aii-Bii). For example, dominant, untrustworthy and cold face shape models correlated with fearful, angry, and sad facial expressions. Submissive, trustworthy and warm face shape models correlated positively with happy (except female submissive and trustworthy) and surprised facial expressions. Finally, competent shape models correlated positively with surprised expressions for female and happy expressions for male faces while incompetent models were mostly associated with sad facial expressions.

Together, these results show that facial expression models of social traits form two clusters that are closely associated with specific emotional facial expressions, in particular happy and disgusted. However, an ablation analysis and correlations between face shape and facial expression models revealed that emotional facial expression features did not fully account for social trait features. Therefore, while social trait and emotion models shared some features, emotion perception alone does not fully account for the perception of social traits.

4.5 Discussion

In this chapter I addressed several key questions regarding which facial expression patterns elicit social trait judgments and how these relate to emotional facial expressions and neutral face shapes of social traits. I did so by applying methods analogous to those in chapter 2 and chapter 3 to facial expressions. Specifically, I used reverse correlation to model the specific Action Unit patterns that elicit social trait perceptions of four key social traits – dominance, competence, trustworthiness, and warmth. This revealed that facial expressions of social traits form two main clusters. The first of these clusters, which included submissive, competent, trustworthy, and warm, was associated with facial movements of the eyebrows and mouth, in particular brow raising and AUs involved in smiles. In contrast, the second cluster of dominant, incompetent, untrustworthy, and cold was associated with furrowed brows and a wrinkling of the nose. Importantly, these two clusters were closely associated with specific emotional facial expressions based on their AU activation patterns, namely happy and sad for the first cluster and disgusted and angry for the second. However, in each case, the social trait facial expression models also comprised AUs that were not associated with the respective emotion. Furthermore, I showed that these social trait facial expressions correlated positively with their respective social trait face shapes with the exception of competence. Comparison of the social trait shape features with Action Unit features further revealed the specific facial features shared between neutral shapes and different facial expressions. For example, trustworthy and warm face shape models comprised raised brows and lip corners similar to facial expression features of happy and surprised.

These results stand in contrast to those of chapter 2 that found that social trait face shape models cluster into four subgroups. Instead, social trait facial expression models grouped into two clusters, aligning with extant work showing that social trait judgments from facial expressions are less variable than those from neutral face shapes (Hehman et al., 2017). Moreover, these two clusters relate to affective facial expressions – happy and disgusted – which may reflect approach and avoidance signals (Adams et al., 2015; McArthur & Baron, 1983) that play a central role in social trait perception (A. L. Jones & Kramer, 2021). Indeed, social trait impressions from neutral faces have been argued to be a substitute for approach and avoidance inferences in the absence of dynamic facial expressions that normally afford such inferences (Todorov, 2008). The results in this chapter support this hypothesis by showing the close alignment between social trait inferences from facial expressions and approach and avoidance signals.

I further found that facial expressions judged as competent comprised features similar to those of submissive, trustworthy, and warm facial expressions while facial expressions judged as incompetent showed higher similarity with dominant, untrustworthy, and cold facial expressions. This is in contrast to perceptions from static face shapes where competent face models share more features with dominant than submissive face models while incompetent face models share features with submissive face models, particularly for male faces (Judd et al., 2005; Oliveira et al., 2019; Sutherland et al., 2016, see also chapter 2). Such divergence in the perception of competence between male and female faces and from face shapes and facial expressions is in line with research showing that judgments of competence tend to be more variable (Hehman et al., 2017; Sutherland et al., 2018) and depend, more so than other social traits, on perceiver characteristics (Hehman et al., 2017) and contextual cues (e.g., Chan-Serafin et al., 2019; Oh et al., 2019). Here, I extend these findings by showing that neutral face features and facial expression features that elicit competence perceptions do not overlap (with the exception of male competent). In other words, competence is perceived from a wider range of features than other social traits, going some way to explain variability in its judgments. Moreover, competence is generally associated with positive social attributes. For example, competence is associated with emotional stability, conscientiousness (Abele et al., 2008; Abele et al., 2016), likability (Willis & Todorov, 2006; Zebrowitz et al., 2010), and perceived leadership ability (Chen et al., 2014; Klofstad, 2017; Sussman et al., 2013). The similarity between competent, warm and trustworthy facial expression models in the current results may reflect such positive associations.

The results further showed that, although facial expression models of submissive, competent, trustworthy, and warm shared AUs with happy facial expressions, they also comprised Eyebrow Raiser and, for female faces, Eyes Closed. Although many of the AUs involved in these social trait facial expressions, specifically Lip Corner Puller, AU25-12 (Lips Part + Lip Corner Puller), and Dimpler, are associated with smiles (Ambadar et al., 2009; Ekman et al., 1990; Frank et al., 1993), it is much less clear whether these form *happy* smiles in combination with the remaining AUs. People smile for many reasons other than happiness, such as embarrassment (e.g., Ambadar et al., 2009; Keltner, 1995; Keltner et al., 1997), to signal appeasement (Goldenthal et al., 1981; Hess et al., 2002; Keltner et al., 1997), affiliation, dominance (e.g., J. Martin et al., 2017; Rychlowska et al., 2017), and physical or psychological pain (e.g., Harris & Alvarado, 2005; Kunz et al., 2009; Singh & Manjaly, 2021) and observers are sensitive to these differences (Calvo et al., 2013; Dibeklioğlu et al., 2012; Krumhuber & Manstead, 2009). For example, eyebrow raising as observed here in social traits is very uncommon in amused smiles but highly common in sentimental ones (McDuff, 2016). In contrast, AU12-6 – a smile with raised cheeks (i.e., Duchenne smile Williams et al., 2001) – as observed in happy but not the social trait facial expressions, is more common in amusement (i.e., happy) smiles than in other types of smiles (Ambadar et al., 2009). Furthermore, smiles can convey lower social status (Ketelaar et al., 2012) and lower physical dominance (Kraus et al., 2013). Similarly, steady and direct eye gaze is a sign of dominance in primates (Bolwig, 1978; Jacobus & Loy, 1981) and humans (Main et al., 2009; Toscano et al., 2018) while lowered gaze is associated with avoidance behaviors (Adams & Kleck, 2005; Jack et al., 2016) such as pain expressions (Prkachin, 2009; Prkachin & Solomon, 2008) and disgust (Fischer et al., 2012; Jack et al., 2016). The closing of the eyes in female competent, trustworthy, and warm facial expression models may therefore reflect such avoidance, submission (Terburg et al., 2011) or appeasement.

These findings extend to social trait shapes such that submissive, competent, trustworthy, and warm face shape models physically resembled AUs involved in happy facial expressions (e.g., Lip Corner Puller). However, they additionally resembled Eyebrow Raiser, Mouth Stretch and Jaw Drop (likely reflecting the lengthening of the face), making them similar to their respective social trait facial expressions (with the exception of competence) but less so to happy expressions. This suggests that any physical similarities between these social trait shape models and happy facial expression models are not necessarily evidence for overgeneralization from positive emotional signals but may be associated with general approach signals (A. L. Jones & Kramer, 2021; Slepian et al., 2017; Todorov, 2008). In contrast, dominant, untrustworthy, and cold face shape models exhibited physical similarity to fearful, angry, and sad facial expressions, suggesting an overgeneralization from negative affective signals (Adams et al., 2012; Montepare & Dobish, 2003). However, similarities between these social traits and fearful and sad expressions appear counter-intuitive and run counter to extant work that showed negative correlations between, e.g., dominant and fearful/sad (Adams et al., 2012; Albohn & Adams, 2021). The similarities here arose from correlations with individual AUs that each form part of several different emotional facial expressions. Ultimately, it is the *combination* of different AUs that conveys a given meaning to the perceiver (Ekman & Friesen, 1978; Jack & Schyns, 2017). In other words, though dominant face models may comprise features also found in sad facial expressions, this does not necessarily equate to dominant faces being perceived as sad. Equally, however, as seen with the happy facial expressions, obvious featural overlaps, such as smiles, can be misleading and may result in oversimplified conclusions (e.g., trustworthy = happy). The current results therefore highlight the difficulty in making inferences based on whole face correlations.

In summary, chapter 4 identified a latent space of social trait facial expressions comprising two clusters that closely aligned with approach and avoidance signals. Moreover, I found only limited evidence for the overgeneralization hypothesis. Specifically, facial expressions eliciting judgments of submissive, competent, trustworthy, and warm were related to appeasement smiles and face shapes of these social traits similarly reflected general approach signals. Results therefore formed evidence for the centrality of approach and avoidance in social trait perception while questioning the wider applicability of emotion overgeneralization accounts.

Chapter 5

General Discussion

5.1 Summary of main findings and contributions

Across three experiments, this thesis examined the face features that elicit perceptions of four key social trait dimensions – dominance, competence, trustworthiness, and warmth – in three separate dimensions – 3D shape, 2D complexion, and dynamic facial expressions. Results revealed that social traits cluster into four subgroups on the basis of their face shape and facial complexion (chapter 2). These subgroups closely align with intent and ability subgroups (Fiske et al., 2002; Fiske et al., 2007; Walker & Vetter, 2016), supporting earlier work showing the centrality of these dimensions (Fiske et al., 2002; Oosterhof & Todorov, 2008). However, chapter 2 extended upon this work by (1) modeling the precise face shape and complexion features that subtend each social trait cluster, (2) thereby pinpointing sources of perceptual variance between the dominance and competence dimensions (Oliveira et al., 2019; Sutherland et al., 2016), and (3) revealing that judgments grouped by intent and ability arise from an underlying latent face feature space.

These social trait face features were furthermore systematically related to other key social judgments and dimensions. For example, social trait face complexion was highly related to age complexion, such that dominant, untrustworthy, and cold face complexions correlated with older aged complexions whereas submissive, trustworthy, and warm face models correlated with younger aged complexions. These results highlight the importance of disentangling the effects of different feature spaces on social trait perception and help explain why some research has found a third, age related, dimension (Sutherland et al., 2018; Sutherland et al., 2013; Vernon et al., 2014). Similarly, social class, a central hierarchy in human societies (Kraus et al., 2013), was perceptually related to stereotypically associated social traits and in particular to competence (chapter 3). However, no single social trait's face features fully accounted for social class features, suggesting social class perception cannot be reduced to these stereotypical associations.

In contrast to social trait face shape, facial expressions that elicit social trait perceptions formed two, not four, distinct clusters and these were likely related to approach and avoidance (chapter 4). Facial expression patterns within these clusters comprised Action Units from stereotypically related emotional facial expressions (e.g., happy) but also additional Action Units. Similarly, social trait face shape features could not be fully accounted for by any single emotional facial expression. These results lend support to the overgeneralization hypothesis only in so far as some social traits' features were associated with generally negative or positive affective signals (e.g., dominant with negative valence). However, in light of evidence that facial expressions of happiness can bear similarity to facial expressions of less positively valenced facial expressions such as pain or embarrassment (Keltner, 1995; Keltner et al., 1997; Kunz et al., 2009; Singh & Manjaly, 2021) and these bear some resemblance to social trait features (e.g., raised eyebrows and closed eyes), it is possible that social trait inferences are either overgeneralizations from other facial expressions than previously hypothesized or are not emotion overgeneralizations at all.

Together, this set of results builds on and expands work showing the interrelated nature of social trait perceptions and perceptions of other socially relevant categories from faces. These social judgments are fundamental to human social interactions (e.g., Chua & Freeman, 2022; Connolly et al., 2021; Fiske et al., 2002; Qi et al., 2021) and decision making (e.g., Antonakis & Eubanks, 2017; Jaeger et al., 2019; Todorov et al., 2005). Given also recent advancements in machine learning to automatically predict humans' stable traits based on such facial cues (Al Moubayed et al., 2014; Keles et al., 2021; S. Song et al., 2021) and the use of such tools in societally relevant ways (e.g., job applicant screening; Bekhouche et al., 2017; Nguyen et al., 2013), understanding how *humans* make these judgments from faces and how these judgments relate to each other becomes crucial for the design of social interventions and ethical software and machines. When designing such interventions, however, one should take into account that although featural correlations between social traits and other social perceptions (e.g., happy) may drive correlations between judgments, these do not necessarily reflect a ground-truth (Todorov et al., 2015; Zebrowitz, 2017), meaning their automatic inference from facial features is ill-advised (Keles et al., 2021).

5.2 Limitations and future directions

5.2.1 Social traits and age perception

The results showed that age-related facial cues, such as color contrast around the eyebrows and mouth, contributed to social trait perception. For example, more dominant faces also comprised cues of older age. This suggests that social trait perception changes with the age of the target face. However, the age of both the stimuli and the participants in the work presented here were constrained to be between 18 and 35 years. It is likely that widening this age range would change the results. This is because people generally favor self-resembling faces (Bailenson et al., 2008; DeBruine, 2005; Krupp et al., 2008; Zebrowitz et al., 2007), as predicted by the familiar face overgeneralization hypothesis (Zebrowitz et al., 2007). With regards to age, people exhibit own-age biases in the perception of hostility, competence, trustworthiness, aggression, and attractiveness (Zebrowitz & Franklin, 2014; Zebrowitz et al., 2013). There is also more in-group than between-group agreement in social trait judgments from younger and older adults (Zebrowitz et al., 2013). And while in the current participant group of 18-35 year olds, competence was not associated with babyfaced features, previous research found older aged adults perceived more babyfaced individuals as more competent (Zebrowitz & Franklin, 2014). Similarly, there is evidence that emotion overgeneralization is greater for older adult faces than younger adult ones (Barber et al., 2019). Together, this suggests that the fundamental features driving social trait perception may change across the life span of perceivers.

Additionally, the current results suggest there are gender differences in the relationship between social trait perception and age. Specifically, I found that for female but not for male faces, trustworthy and warm complexions also comprised older age cues, in particular lower contrast (PC3). Modeling the precise facial features of older and younger faces as judged by older and younger perceivers on central social traits would yield further insight into these relationships. Considering that people show preferences for self-resembling faces in important contexts such as cooperation for the public good (Krupp et al., 2008) and voting choices (Bailenson et al., 2008), understanding these perceptual differences seems particularly urgent.

5.2.2 Cross-cultural differences in social trait and social class perception

Furthermore, there is clear evidence for cross-cultural differences in social trait perception (B. C. Jones et al., 2021; Sutherland et al., 2018). For example, age perception may be more relevant than competence or dominance perceptions in Asian cultures (Sutherland et al., 2018). Historically, elder respect has been central in Asian, collectivist societies (Sung, 2001, 2004), perhaps explaining the centrality of age in social perception in these cultures. However, industrialization and aging populations are negatively impacting positive attitudes towards elders (e.g., North & Fiske, 2015). It remains an open question how such beliefs impact perceptions of social traits generally, and across the life-span.

Similarly, future work should further examine whether and how conceptual knowledge of and beliefs about social class might impact the facial features that drive these perceptions, particularly in different cultures where social class stereotypes (Grigoryan et al., 2020; Schofield et al., 2021) or physical attributes associated with health or beauty might vary (e.g., Zhan et al., 2021). Investigating the conceptual and perceptual similarities between social traits and social class in different cultures might yield further insight into the complex association between social class and social trait perception. Ethnicity but also age-related facial features furthermore correlate with social hierarchies and are involved in stereotype knowledge (e.g., Brown-Iannuzzi et al., 2017; Freeman & Ambady, 2011; Lei & Bodenhausen, 2017) and impact on first impressions (Xie et al., 2019; Xie et al., 2021), suggesting they might relate to both social class and social trait perception.

5.2.3 Do femininity and masculinity features correspond to central social trait features

Another fruitful avenue of investigation will lie in identifying the specific facial features that drive masculinity and femininity perceptions. A recent conceptual shift suggests that the concepts of masculinity and femininity may be the two primary dimension of social trait perception (A. E. Martin & Slepian, 2021). Perceived gender also impacts a wide range of first impressions. For example, female faces are perceived to be less competent than male faces (Oh et al., 2019) and masculine and dominant-looking female faces are judged negatively (Sutherland et al., 2015). Facial femininity and masculinity have also been shown to be predictors of trustworthiness and dominance judgments, respectively (Jaeger & Jones, 2022). However, the current work did not measure perceptions of femininity and masculinity and sexual dimorphism (an objective measure) may be only weakly related to masculinity and femininity judgments (Hester et al., 2021). An investigation of the relationship between perceived masculine and feminine face features and the fundamental social trait feature spaces identified here may thus yield further insight into the importance of gender-perceptions for first impressions. Indeed, reverse correlation is well suited to first modeling facial features that elicit masculinity and femininity judgments and then measure their similarity to other key social dimensions.

5.2.4 Facial expressions that elicit social trait judgments

Additionally, though chapter 4 is, to date, the most comprehensive investigation of the association between social trait perception and facial expressions, many questions remain unanswered. For example, different types of smiles (e.g., happy, embarrassed, etc.) differ not only in their AU patterns but also in the temporal activations of these AUs (Ambadar et al., 2009). Such differences are likely to also emerge in AUs shared between emotional facial expressions and social traits. Though chapter 4 did model temporal parameters, a comparison of these between the social traits and emotional expressions was prevented by the differing study designs of the two experiments, making effects caused by experimental noise more likely. It similarly remains an open question whether smiles associated with submissive, competent, trustworthy, and warm resemble any particular smile type other than happy. To test this, one would need to derive facial expression models of potentially relevant smiles such as embarrassment or appeasement and compare their specific AU patterns to those of each social trait. Similarly, it remains to be tested empirically whether the two social trait facial expression clusters identified here correspond to approach and avoidance.

An additional factor potentially influencing the perception of social traits from facial expressions is face complexion. People perceive emotions from face color alone (Thorstenson, 2018; Thorstenson et al., 2018). For example, green and blue face complexions elicit perceptions of sadness and fear while red is particularly associated with anger but also with happiness (Thorstenson, 2018). These perceptual associations have a physiological basis. Specifically, high arousal emotions such as anger, happiness, and surprise involve a reddening of the face due to a widening of the blood vessels (Drummond, 1994; Levenson, 2003; Shimbo et al., 2013). In contrast, low arousal emotions, including sadness, fear, and disgust involve a constriction of blood vessels – blanching (Hayashi et al., 2009; Kreibig et al., 2007; Rohrmann & Hopp, 2008). Considering the tight association between social trait perception and affective signals (Knutson, 1996; Said et al., 2009; Zebrowitz, 2017), it is possible that such facial coloration interacts with facial expressions to elicit certain social trait perceptions, though the impact of facial expressions is likely to be greater due to the greater salience of the movement. One possibility is that that facial coloration that is congruent with the perceived emotional valence of a given social trait could intensify the perceived social trait (i.e., Nose Wrinkler with red face = extremely dominant). Alternatively, face color may facilitate social trait perception such that congruent face complexion (e.g., red face and angry expression) leads to faster judgments of the associated social trait (e.g., dominant) than incongruent complexion. These are speculations that remain to be tested although there is some converging evidence to support them. For example, smiling expressions increase perceived skin lightness (H. Song et al., 2012), which aligns with generally lighter complexions of trustworthy and warm social trait models (chapter 2). Additionally, while face shape can be transformed through muscle movement (i.e., facial expressions) to resemble a certain social trait (Gill et al., 2014), overall face complexion does not physically change with facial expressions, making an interaction between complexion and facial expressions more likely as the complexion's effects cannot be masked. In contrast, details, such as wrinkles, *can* be enhanced by facial expressions. For example, Nose Wrinkler produces folds around the nose (Ekman & Friesen, 1976; Y. Zhang & Ji, 2005) and such folds are related to the perception of dominant, untrustworthy, and cold (chapter 2), suggesting at least some featural overlap between social trait and emotion complexion. However, as the current work did not measure face coloration in relation to emotion perception, these questions remain unanswered.

5.2.5 Imbuing virtual agents with task-relevant social traits

More broadly, the results presented in this thesis can inform the design of humanoid artificial agents and social robots. While robots in general do not require a humanoid appearance, humanlike facial appearance can improve perceptions of robots and agents, depending on the task (Broadbent et al., 2013; Duffy, 2003; Prakash & Rogers, 2016). Imbuing such social agents with different social traits is particularly important in fields where users are vulnerable and trust towards the virtual agent/robot is crucial. Examples include virtual therapists and support agents (Brander et al., 2021; DeVault et al., 2014; Marsella & Gratch, 2003; Philip et al., 2017; Rizzo et al., 2011), medical (Farrier et al., 2020; Tan et al., 2022) and inter-personal training (P. Kenny et al., 2007; Traum et al., 2005), and education (e.g., anti-bullying Aylett et al., 2009). In each of these situations, robots may benefit from a warm, trustworthy, and/or competent appearance. Perceived social traits of such agents also impact interaction enjoyment (Lee et al., 2006) and likability (Zhong et al., 2022). Increasing perceived trustworthiness can furthermore reduce perceived 'creepiness' of artificial agents (Watt et al., 2017). However, imbuing such agents with social traits requires two things: First, the facial features found here need to be transferred to virtual agents and social robots. This may not be trivial as, for example, artificial faces receive lower trustworthiness ratings than real faces, although this may be particularly due to differences in the eyes (Balas & Pacella, 2017). As the stimuli in this thesis were photorealistic but the features derived from them may be transferred to entities with less humanlike appearance, it remains an open question whether these features still elicit the desired social traits. Second, a greater understanding of what drives perceptions of relevant social traits in these virtual agents and social robots is needed. The current thesis goes some way in building a comprehensive understanding of the types of *facial* features that elicit positive social perceptions. However, to use them for the development of such agents, these need to be integrated with other modalities, such as gestures and voice. In sum, imbuing social agents and robots with task-relevant social traits via their facial features as derived through psychological methods can benefit many important social applications but requires further research for successful implementation.

5.3 Concluding remarks

In conclusion, this thesis revealed the latent 3D shape, 2D complexion and dynamic facial expression features that elicit perceptions of four key social trait dimensions – dominance, competence, trustworthiness, and warmth. These results support previous work showing that two orthogonal dimensions, namely dominance/competence/ability and trustworthiness/warmth/intent, shape human social inferences and that social traits share features with social class and with affective signals that may reflect approach and avoidance attributions. At the same time, current results challenge ideas about the overgeneralization of emotional facial expressions to social inferences. In sum, these results inform theories of social perception and can furthermore guide the design of digital social agents although it should be considered in these applications that such social trait inferences seldom reflect ground-truths (e.g., Todorov et al., 2015).

Chapter 6

Supplementary material

6.1 Participant questionnaire

Each participant completed the following questionnaire to assess exposure to non-Western cultures prior to being admitted to the study. I included only those participants who answered 'no' to all questions or, if they answered 'yes' to question 2, had only been on vacation in non-Western countries for a short period of time (at most three weeks) or not in the recent past (at least two years).

1. Have you ever lived in a non-Western* country before (e.g., on a gap-year, summer work, move due to parental employment)?
2. Have you ever visited a non-Western country (e.g., vacation)?
3. Have you ever dated or had a very close relationship with a non-Western person?
4. Have you ever been involved with any non-Western culture societies/groups?

*By Western groups/countries, we are referring to Europe (Eastern and Western), USA, Canada, United Kingdom, Australia, and New Zealand

6.2 Determining required number of trials – 3D Shape and 2D Complexion

A previous experiment investigating the related social attribute attractiveness (Boshyan et al., 2014; Jaeger et al., 2020) using reverse correlation based on larger trial numbers (n = 1,950; Zhan et al., 2021) had indicated that fewer trials ($\sim$1,200) were sufficient to achieve stable model estimates. To confirm that 1,200 trials were sufficient to similarly derive stable estimates of 3D shape and 2D complexion models for each social trait, I collected pilot data for four participants (white Western; 2 female, 2 male; mean age = 23.75 years, *SD* = 4.79 years) with procedures as described in chapter 2. Specifically, each participant rated 1,200 random male face identities on each one of four social traits (dominance, competence, trustworthiness, warmth) on a seven-point scale from, e.g., 'very submissive' to 'very dominant' with 'don't know' as the central fourth button. Face identities were generated using an earlier version of the GMF based on a reduced set of identities (350), resulting in fewer identity components (350 for shape and 350 $\times$ 5 for complexion).

Following the experiment, I derived shape and complexion models for each social trait and participant using ridge regression. Specifically I regressed the social trait judgments (n = 1,200 per social trait) onto the individual identity components (350 shape and 350 $\times$ 5 complexion) of stimuli presented at each trial. Next, I repeated the model building procedure (i.e., ridge regression) 120 times based on subsamples of trials. Specifically, for each social trait and participant, I derived 120 separate shape and complexion models based on increasing numbers of trials (10-1,200 in steps of 10). At each iteration, I estimated bootstrapped 95% confidence intervals (100 bootstrap samples) of the beta coefficients. Figure 6.1 show the resulting 95% confidence intervals (z-score; mean across participants and identity components) for shape (left) and complexion (right) with estimated beta coefficients (z-scores) on the y-axis and number of trials on the x-axis. The figure suggests that confidence intervals do not change drastically after around 400 trials. However, considering the high-dimensionality of the feature space in consideration, I chose to retain 1,200 trials per social trait and sex of stimulus face in the final experiment.

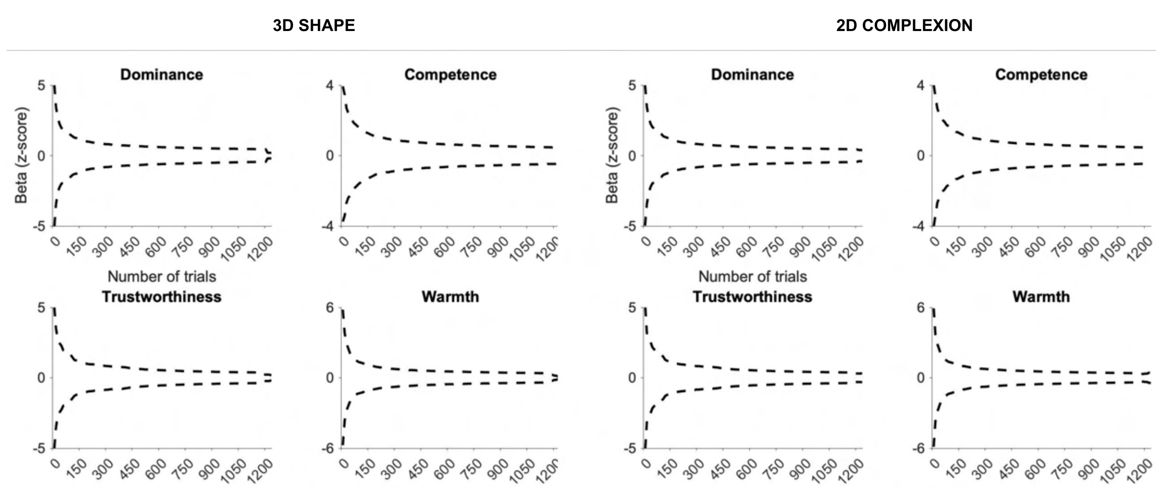

Figure 6.1. Confidence intervals (95%) of 3D shape and 2D complexion model beta coefficients. Each plot shows bootstrapped (100 samples) 95% confidence intervals of the model beta coefficients for each social trait (see plot titles) and shape (left) and complexion (right) separately at increasing numbers of trials (x-axis). Beta coefficients (y-axis) are z-scored for comparability and shown as mean across identity components and participants.

6.3 Testing the linearity assumption for modeling 3D shape and 2D complexion

Using a linear regression to model the shape and complexion that elicit social judgments would be unwarranted should the relationship between either of these modalities and social judgments be non-linear. To test the linearity assumption I proceeded in the following steps. First, for each sex of stimulus face, social judgment (dominance, competence, trustworthiness, warmth, and social class) and shape and complexion separately, I computed the average (median) across face stimuli for each one of four rating bins. Figure 6.4 through Figure 6.5 (right) show these face models for the lowest and highest rating bin and for each sex of face, social judgment, and shape and complexion separately. Next, for each average model, I quantified the deviation from the average face as described in subsection 2.3.7 of chapter 2. This yielded a 14,319 $\times$ 4 matrix of distances for shape [14,319 vertices $\times$ 4 rating bins] and a 5,562 $\times$ 4 matrix of distances for complexion [1,854 downsampled pixels $\times$ 3 L*a*b $\times$ 4 rating bins]. Figure 6.4 through Figure 6.5 show these matrices on the left for each social judgment separately. Blue corresponds to negative deviation from the average face and yellow to positive deviation from the average face. A linear relationship between the deviation from the average face and social judgments should be reflected as a linear transition (i.e., from blue to yellow or yellow to blue) from rating bins 1 to 4 for each vertex and pixel (i.e., per row). Figure 6.4 through Figure 6.5 do largely show this pattern, confirming that the relationship between vertex and pixel deviations from the average face and social judgment ratings was indeed linear.

To further confirm this relationship, I applied a K-means clustering to each of these distance matrices separately. Colored bars next to each matrix show the corresponding clusters (matrices are ordered by these; number of clusters was determined via silhouette scores – maximum mean score across 10 iterations) and line plots show the deviation from the average face at each rating bin (i.e., cluster centroids) for each cluster. Line width corresponds to the number of vertices/pixels within each cluster. These line plots therefore show the patterns of change within each cluster from rating bin 1 (submissive, incompetent, untrustworthy, cold, poor) to rating bin 4 (dominant, competent, trustworthy, warm, rich). The near straight increasing or decreasing lines primarily confirmed the linear relationships observed in the color matrices. Together, these results indicate that the linearity assumption was warranted for modeling the 3D shape and 2D complexion of social trait and social class judgments.

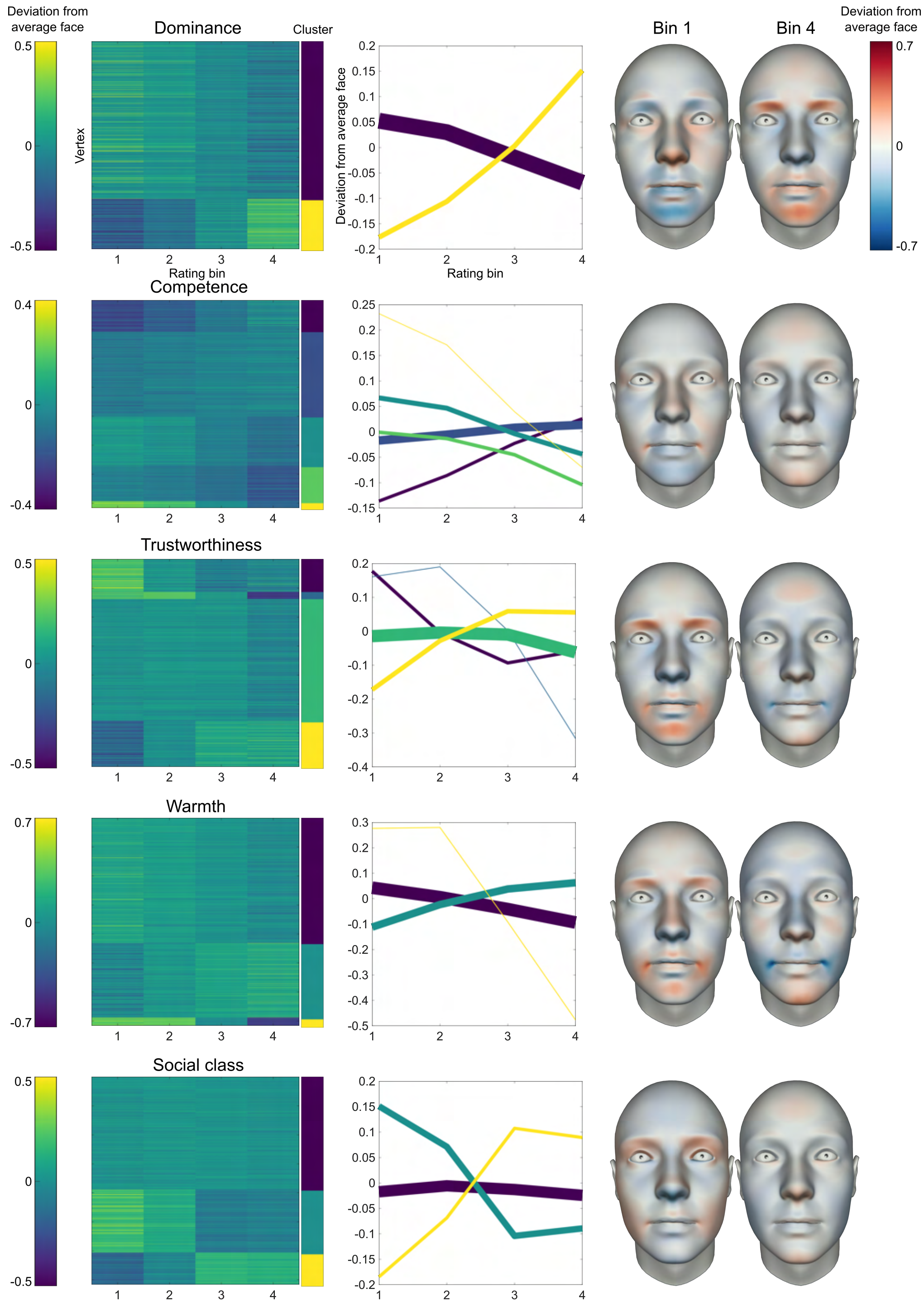

Figure 6.2. Linearity check – female face shape. Colored matrices show each vertex's (n = 14,319) deviation from the average face (blue = negative; yellow = positive; see colorbar to left) in each one of four rating bins (median across participants and trials) for each social judgment. Clusters (K-means) within these matrices are shown to the right. Line plots show cluster centroids at each rating bin. To the right, colored face maps show the median 3D shape features (deviation from the average face; see colorbar to right) for the lowest and highest rating bins (median across participants and trials), normalized across all five social judgments.

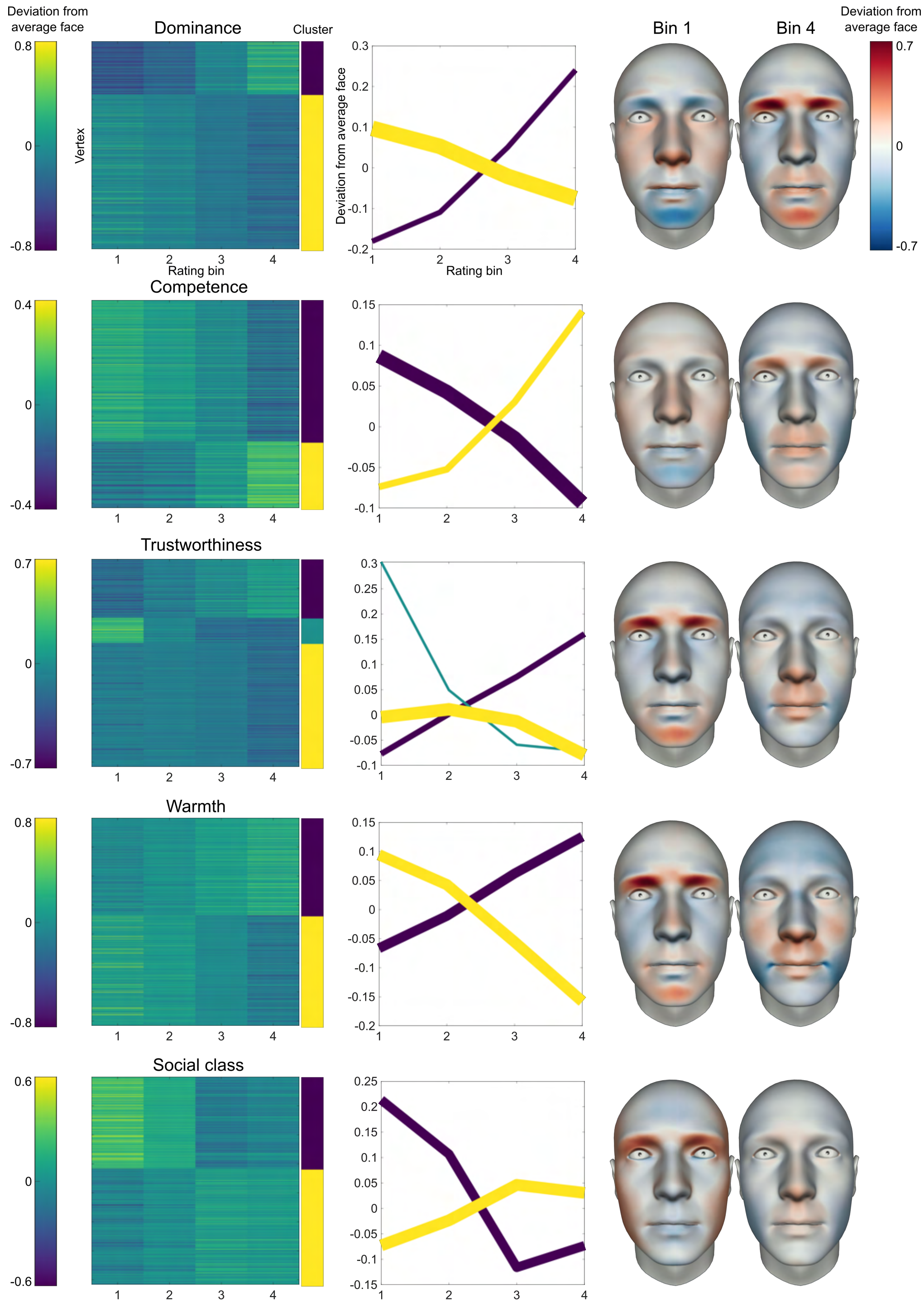

Figure 6.3. Linearity check – male face shape. Colored matrices show each vertex's (n = 14,319) deviation from the average face (blue = negative; yellow = positive; see colorbar to left) in each one of four rating bins (median across participants and trials) for each social judgment. Clusters (K-means) within these matrices are shown to the right. Line plots show cluster centroids at each rating bin. To the right, colored face maps show the median 3D shape features (deviation from the average face; see colorbar to right) for the lowest and highest rating bins (median across participants and trials), normalized across all five social judgments.

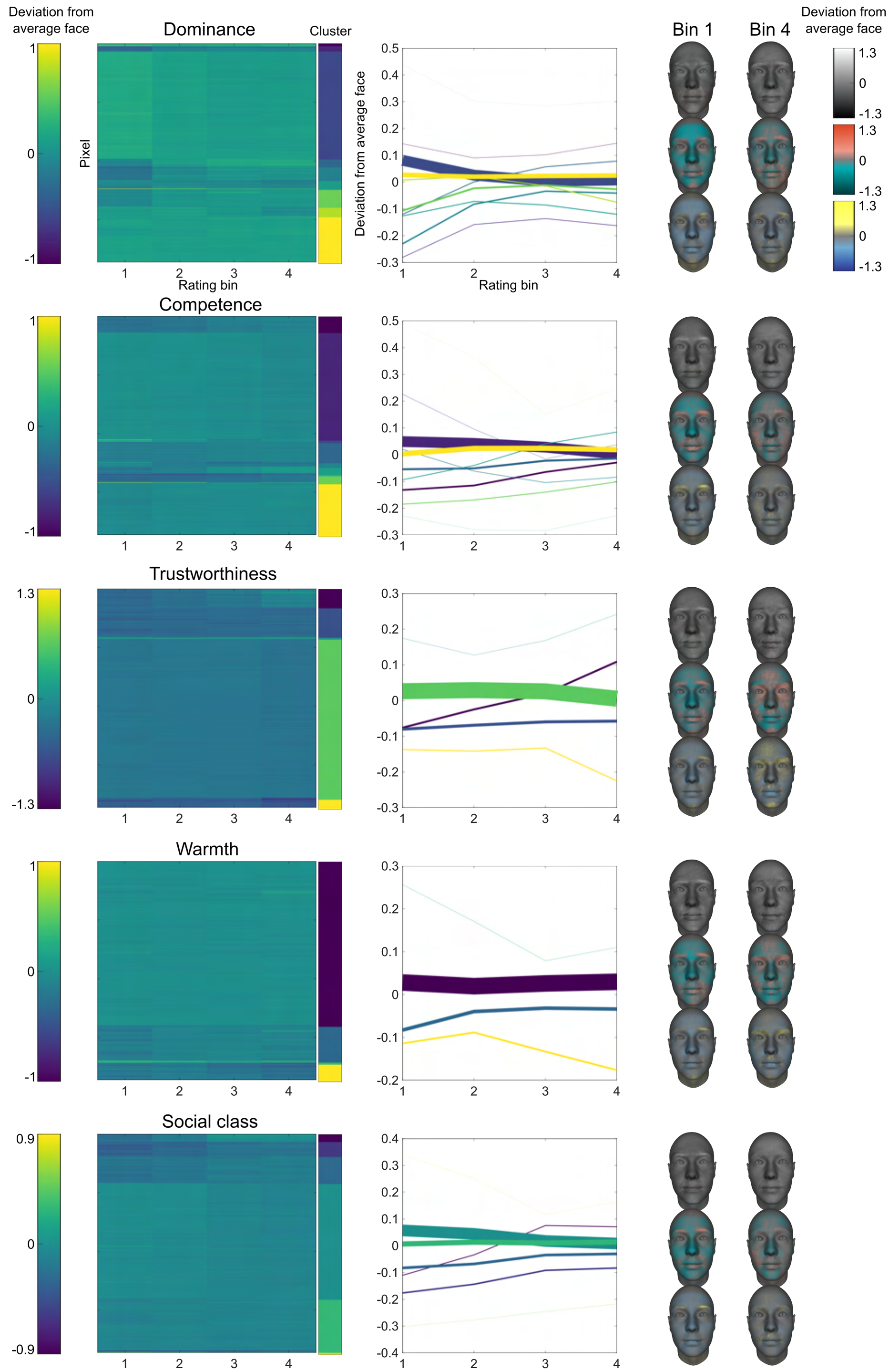

Figure 6.4. Linearity check – female face complexion. Colored matrices show each pixel's (n = 1,854 (downsampled) $\times$ 3 L*a*b) deviation from average face (blue = negative; yellow = positive; see colorbar to left) in each one of four rating bins (median across participants and trials) for each social judgment. Clusters (K-means) within these matrices are shown to the right. Line plots show cluster centroids at each rating bin. To the right, colored face maps show the median 2D complexion features for L*a*b separately (deviation from the average face; see colorbars to right) for the lowest and highest rating bins (median across participants and trials), normalized across social judgments and color channels.

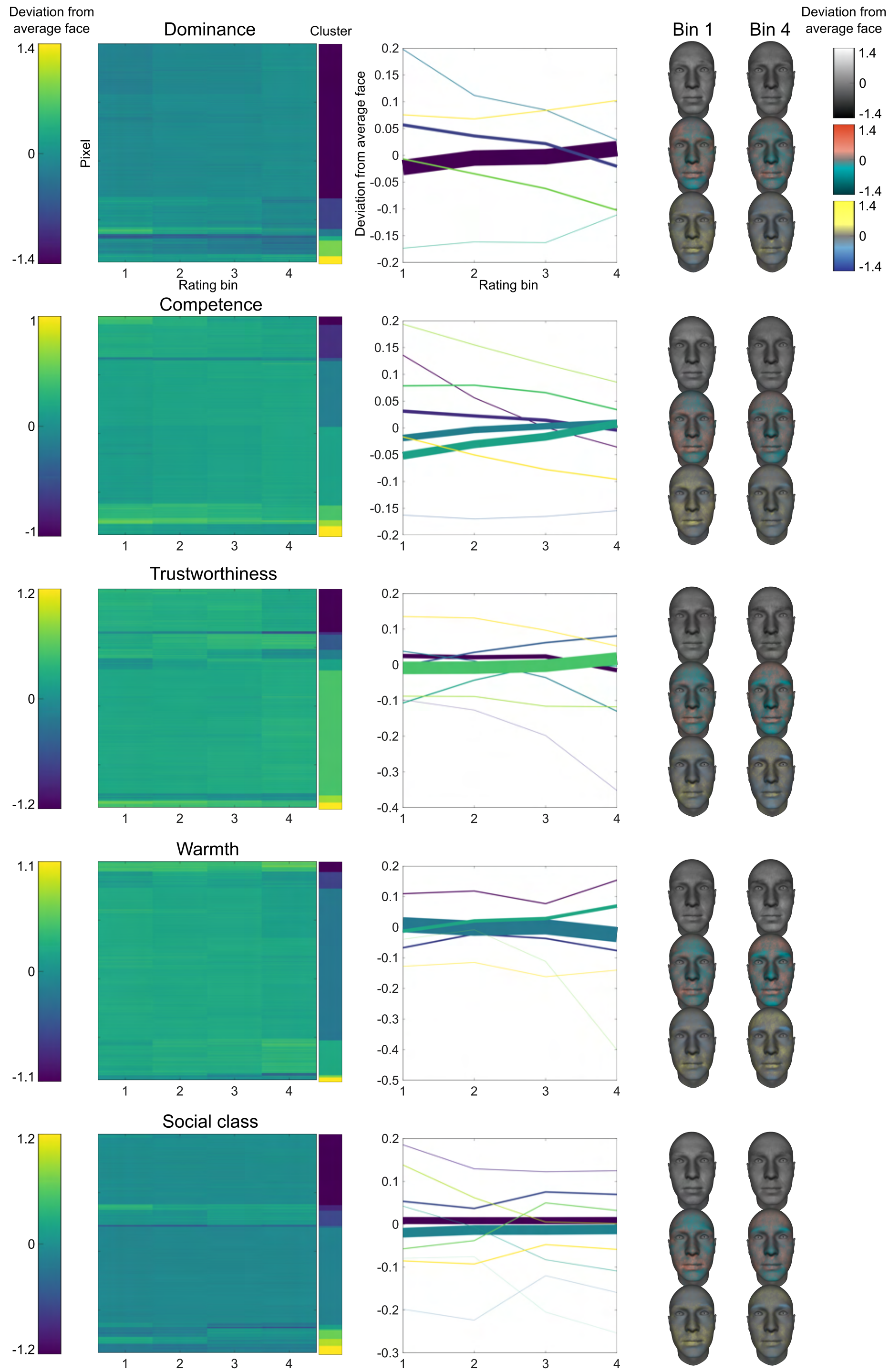

Figure 6.5. Linearity check – male face complexion. Colored matrices show each pixel's (n = 1,854 (downsampled) × 3 L*a*b) deviation from average face (blue = negative; yellow = positive; see colorbar to left) in each one of four rating bins (median across participants and trials) for each social judgment. Clusters (K-means) within these matrices are shown to the right. Line plots show cluster centroids at each rating bin. To the right, colored face maps show the median 2D complexion features for L*a*b separately (deviation from the average face; see colorbars to right) for the lowest and highest rating bins (median across participants and trials), normalized across social judgments and color channels.

6.4 Determining optimal social trait PCs

To identify the optimal number of PCs for social trait shape, complexion and each sex of stimulus sex separately I proceeded as follows. First, as a primary diagnostic tool, I used the elbow method, plotting variance explained by each PC (see Figure 6.6A and Figure 6.7A left). In addition, I plotted the second derivative of the variance accounted for as an indicator of the rate of change between PCs (Figure 6.6A and Figure 6.7A right). Both tools pointed to a 3-component solution for shape and complexion. However, the variance explained by each PC may not only include social trait specific variance, i.e., variance among the social trait models that is associated with social trait variation, but also participant-specific variation. In other words, some amount of variance explained by each PC will be due to idiosyncratic participant differences rather than primarily due to social trait variance. As I aimed to identify facial features associated with social traits, I therefore evaluated the extent to which each of these three PCs was associated with variance across the social traits or variance originating from idiosyncratic participant differences.

To measure this, I first calculated the Mutual Information (MI; Shannon, 1948) between the 229 $\times$ 1 [4 social traits $\times$ 2 poles $\times$ 30 participants] vector of PC scores (220 in the case of complexion) and a vector of the same length of corresponding trait numbers, i.e., 1 (submissive) through 8 (warm). Here, a low MI value would indicate that the PC scores did not vary systematically depending on which social trait category each face shape or complexion model belonged to; a high MI value would indicate that the PC scores varied systematically depending on which social trait category each shape or complexion model belonged to. I then repeated this procedure, replacing the trait numbers by participant numbers, i.e., 1 through 30, associated with each PC score. Here, a low MI value would indicate that the PC scores did not vary systematically depending on which participant each facial shape or complexion model belonged to; a high MI value would indicate that the PC scores varied systematically depending on which participant each shape or complexion model belonged to. I thus quantified the information shared between the PC scores and the social trait category related variance and between the PC scores and the participant-related variance. Finally, I established the statistical significance of each MI value using a Monte Carlo simulation method (one-tailed) with 10,000 iterations.

The resulting distributions are shown in Figure 6.6B and Figure 6.7B for each sex of stimulus face, PC, and shape and complexion separately; solid green lines represent statistically significantly high MI values (p < .05) for participants; solid purple lines represent statistically significantly high MI values (p < .05) for social traits. Non-significant MI scores are indicated with a dashed line. Results showed that, for shape, the PC1 and PC2 scores were statistically significantly more strongly associated with social trait variance than with participant variance. In contrast, PC3 was not significantly associated with social trait

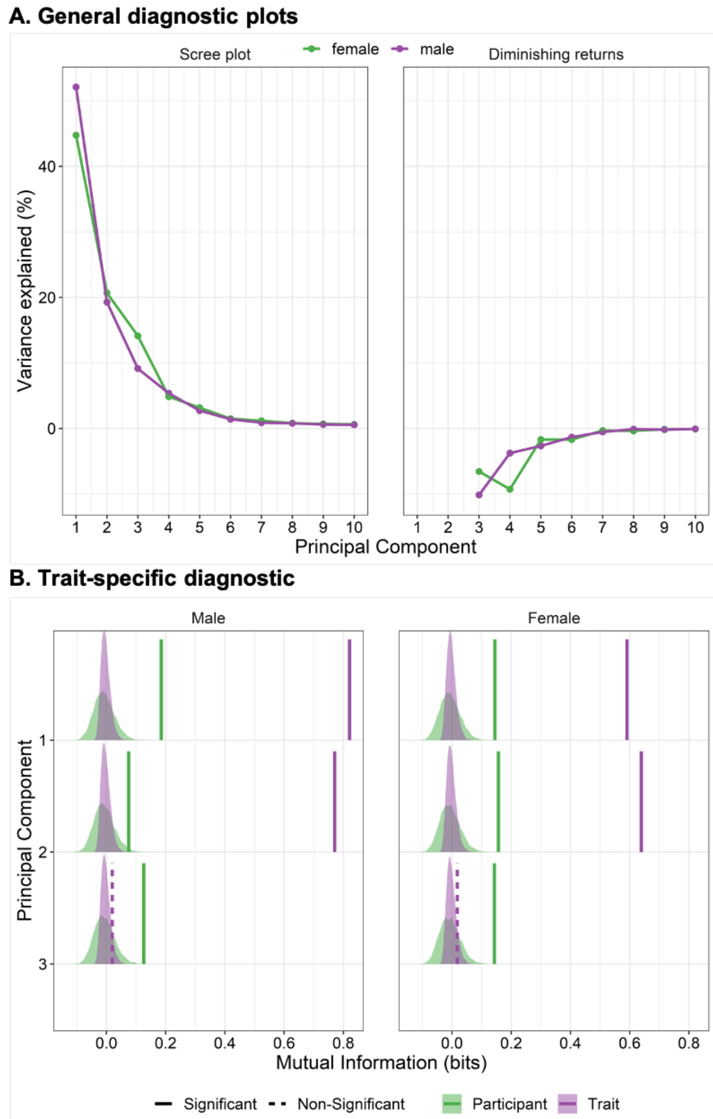

Figure 6.6. Social trait shape choice of principal components. (A) Scree plot shows variance explained by each PC (left). Color here represents sex of stimulus face (green = male, purple = female). In addition, I show the rate of change in the variance explained (second derivative) by each PC (right). (B) To identify which of the PCs are significantly associated with trait-related variance, as opposed to participant-related variance, I calculated Mutual Information between the PCA scores and either participant information (green) or social trait information (purple). Density plots show simulated MI values derived using Monte Carlo simulations. Next, I used a one-tailed test (alpha = .05) to determine statistical significance of the actual MI between each PC and participants/traits, indicated by the vertical lines (solid = significant, dotted = not significant). This indicated that the first two PCs were significantly associated with social trait variance.

variance but was statistically significantly associated with participant variance. Given the primary focus on identifying fundamental social trait face features rather than idiosyncratic differences in social trait perception, I selected a two-factor PC solution to represent the main sources of shape variance across social traits. For complexion, all three PCs were statistically significantly associated with social trait variance suggesting that three PCs were needed to describe social trait complexion features. To confirm this, I included the following additional analysis for complexion only.

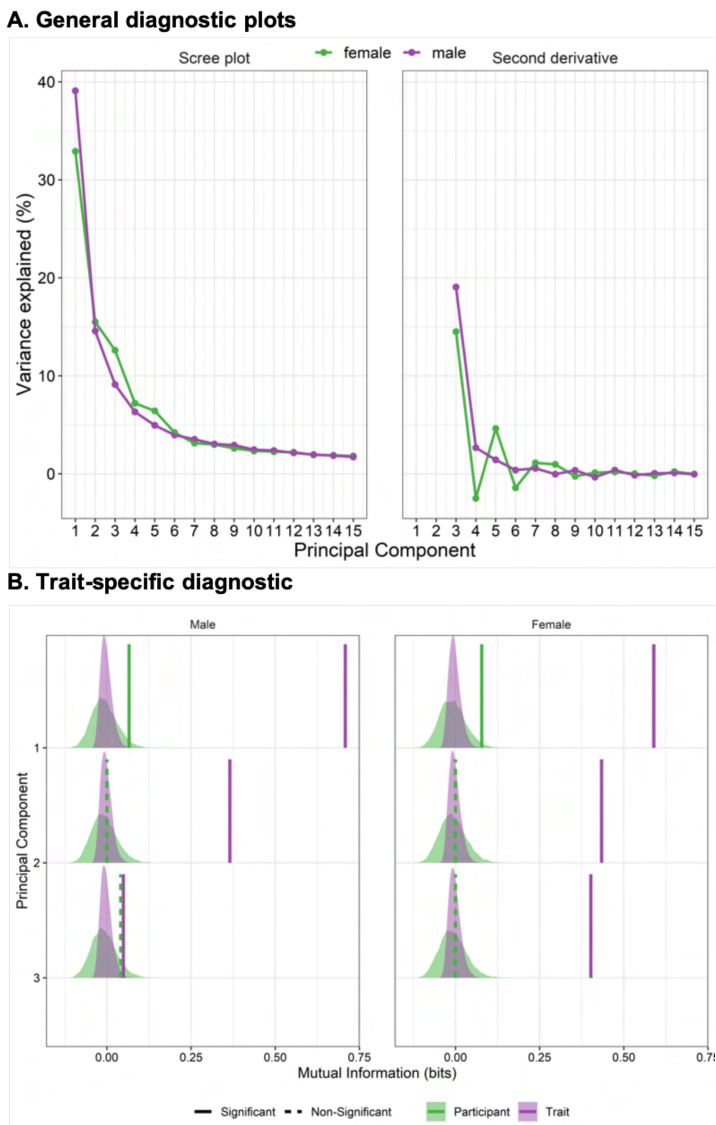

Figure 6.7. Social trait complexion choice of principal components. (A) Scree plot shows variance explained by each PC (left). Color here represents sex of stimulus face (green = male, purple = female). In addition, I show the rate of change in the variance explained (second derivative) by each PC (right). (B) To identify which of the PCs are significantly associated with trait-related variance, as opposed to participant-related variance, I calculated Mutual Information between the PCA scores and either participant information (green) or social trait information (purple). Density plots show simulated MI values derived using Monte Carlo simulations. Next, I used a one-tailed test (alpha = .05) to determine statistical significance of the actual MI between each PC and participants/traits, indicated by the vertical lines (solid = significant, dotted = not significant).

To test how much information each additional PC added in terms of complexion face feature information, I first obtained predicted complexion models, for L*a*b separately, based on a one through five PC solution. For example, the one PC solution model was based on the first PC only, the two PC solution model was based on the first and second PC and so forth. Visual inspection of the resulting faces (see Figure 6.8) indicated that complexion information added by the fourth and fifth PC was not detectible by the human eye. To formally test this, I calculated the rate of change from each PC model to the next as follows. First, I calculated, for each predicted complexion PC model and in each color channel separately, the Euclidean Distance (*d*) between each PC solution and the previous PC solution. For example, for the model based on the first and second PC, I calculated *d* between this model and the model based on only the first PC. As the model based on the first PC only did not have a natural predecessor model, I calculated for this model the Euclidean Distance to the average complexion model. Next, I calculated the rate of change with each added PC (i.e.,

$$dPC_n - dPC_{n-1}/dPC_{n-1}$$

). This revealed that, for both male and female complexion and for each color channel, the rate of change decreased upon adding the fourth PC and thereafter, indicating that these additional components added little additional complexion information. I therefore retained three principal components for social trait complexion.

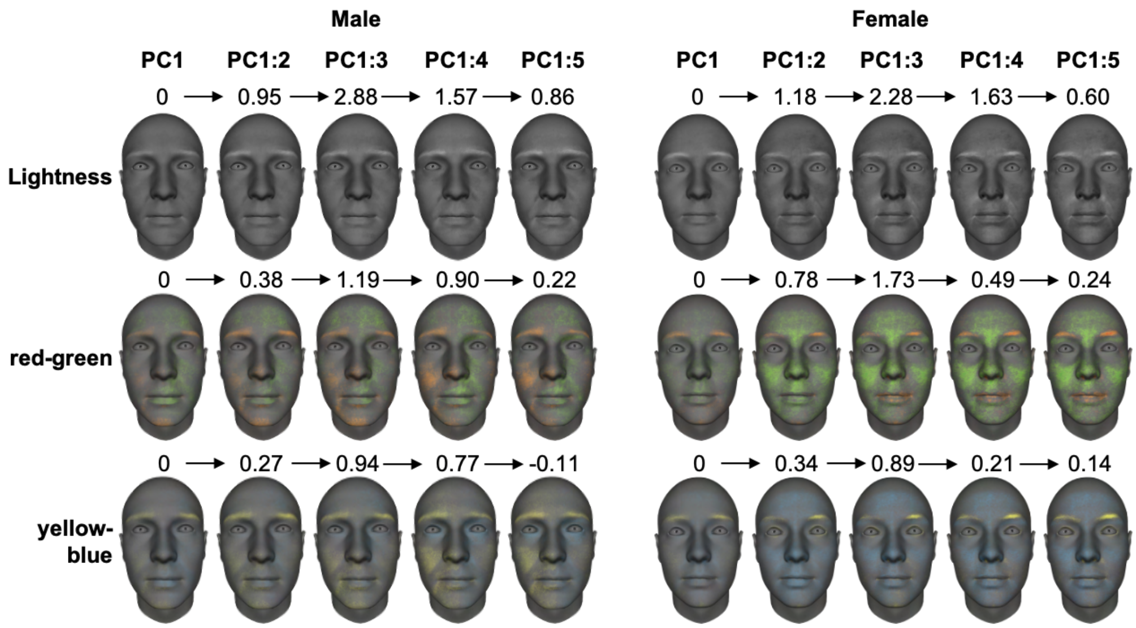

Figure 6.8. Social trait complexion choice of principle components based on added face information. To further clarify the optimal number of PCs, I plotted the predicted complexion effects in L*a*b separately based on 1 to 5 components, successively adding each additional component's face information to the predicted complexion features. In addition, above each predicted model I show the rate of change compared to the previous model (*dPC – dPC-1 / dPC-1*). For example, for male complexion in the Lightness channel, the rate of change increased until the third PC after which it decreased.

6.5 Determining optimal K-means clustering solution

To derive the optimal K-means cluster solution, I computed the K-means solution for 1 through 8 factors, calculating silhouette scores at each iteration. Silhouette scores describe the average distance between each point within a cluster divided by the average distance between all clusters and therefore give a single value per data point describing the goodness of the cluster solution (Rousseeuw, 1987). To select the optimal cluster solution, I first excluded those solutions with any negative silhouette scores, which indicate poor cluster fit, and solutions that included clusters with fewer than 2% of the data points. Next, from the remaining cluster solutions I chose the solution with the highest mean silhouette score. This yielded a two-cluster solution for both female and male PC1 and PC2. Note that when retaining female competence models for PC1 (i.e., models non-significantly associated with PC1) a 3-factor solution best described the results for PC1 with competence forming the third cluster. To plot the cluster solutions, I indexed the data points within each cluster and used the cluster centroid and the mean of the these data points as x and y coordinates, plotting ellipses around them at a distance of 1.5 *SD*.

6.6 Determining required number of trials – Facial expressions

To determine the number of trials required to derive stable facial expression models of social traits, I proceeded similarly as for 2D shape and 3D complexion (see section 6.2). Specifically, I first used previous work to estimate the number of trials needed before confirming this estimate with pilot data. Earlier experiments demonstrated that many facial expressions can be modeled using reverse correlation based on as little as 300 trials (Jack et al., 2014; Jack et al., 2016). However, social trait facial expressions in particular (i.e., trustworthiness and dominance) have been modeled based on 1,200 trials (Gill et al., 2014). To confirm this number, I collected pilot data for five participants (white Western, 3 female; 2 male; mean age = 24 years, *SD* = 5.15 years) with procedures similar to those described in chapter 4 with two exceptions: Firstly, each participant viewed 1,200 random dynamic facial expressions per sex of stimulus face and social trait (i.e., 9,600 unique facial expression stimuli). Secondly, participants rated each dynamic facial expression on a scale from 1 to 6 (e.g., 'very submissive' to 'very dominant') with 'don't know' as the fourth central button.

Following the experiment, I derived facial expression models for each sex of stimulus face, social trait, and participant separately using methods as in chapter 4. Specifically, I obtained significant AUs based on Monte Carlo simulation (one-tailed; 800 permutations). Then, for each significant AU, I used constrained least squares regression to estimate the temporal parameter values. Next, I repeated this procedure, basing each model on a subsample of trials (i.e., 100-1200). Finally, to test model variability with increasing trial numbers, I correlated (Pearson's r) each facial expression model with its predecessor (i.e., the model based on 300 trials with the model based on 200 trials). Figure 6.9 shows the resulting correlations for each sex of stimulus face (panel A) and each participant (panel B), separately. Note that due to an experimental error, not all participants completed 1,200 trials for each social trait dimension. As a result, I could not derive models for each social trait for all participants. Nevertheless, Figure 6.9 demonstrates that after initial high variability, models generally become stable very quickly (see panel A) with high correlations between each model-pair. However, there is some participant variability (e.g., see panel B, participant 5), suggesting that higher trial numbers may still be valuable. Based on this converging evidence, I therefore chose to include 1,000 trials per sex of stimulus face and social trait.

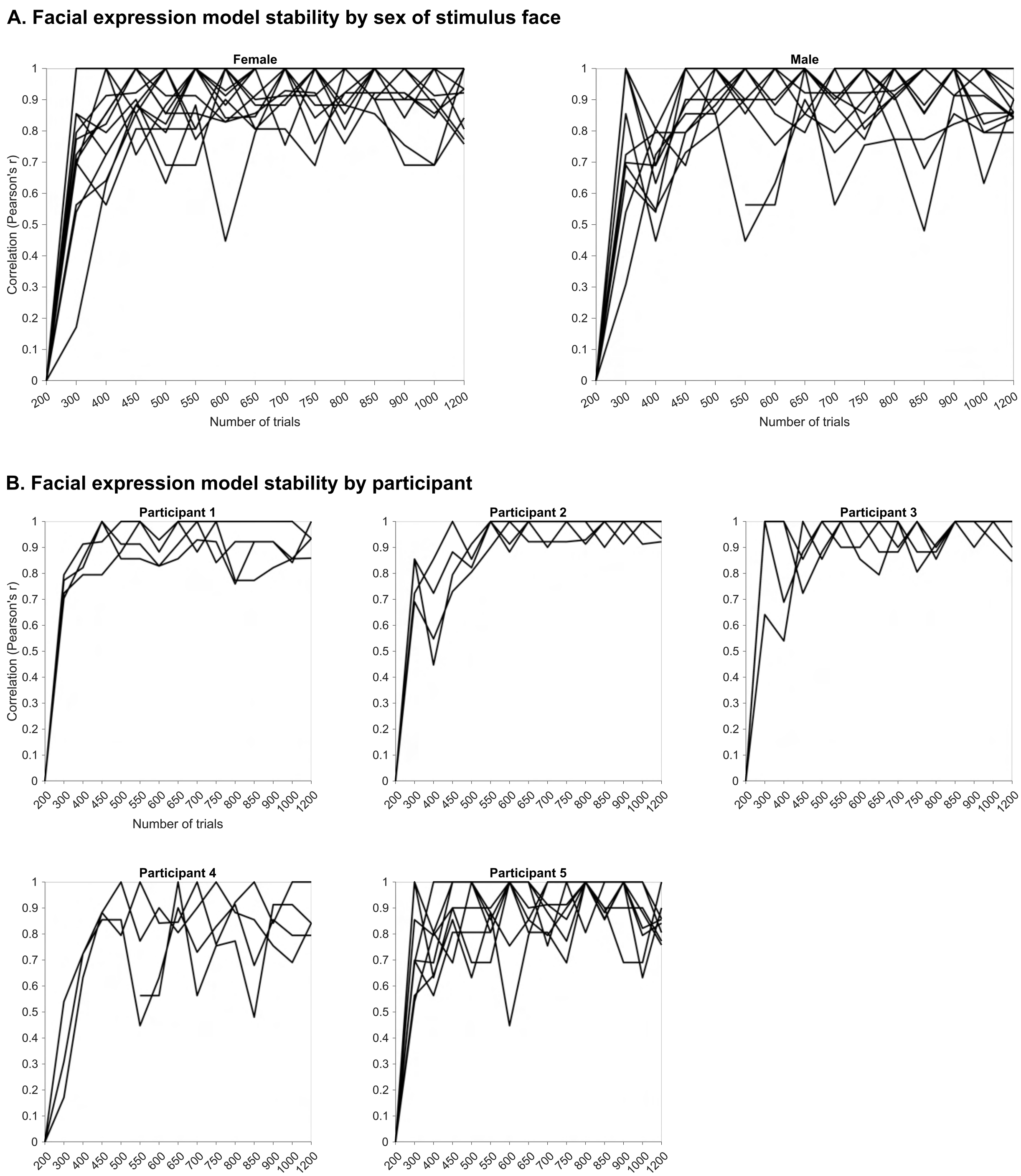

Figure 6.9. Model stability of social trait facial expression models. (A) Facial expression model stability by sex of stimulus face. Each plot shows the correlation (Pearson's r; y-axis) between each facial expression model and its predecessor model for female and male faces. Each line corresponds to one participant and one social trait. Number of trials on the x-axis. (B) Facial expression model stability by participant. Results are shown as in A for each participant separately.

6.7 Deriving emotional facial expression models

The facial expression models of the six basic emotions (happy, fearful, disgusted, angry, sad) were derived in a separate experiment that was not carried out as part of this thesis. For full details see Jack et al. (2014). The experiment derived emotional facial expression models using reverse correlation. Specifically, sixty white Western perceivers (31 female, 29 male; mean age = 22 years, *SD* = 1.71 years) viewed randomly generated dynamic facial expressions (generated using the GFG; Yu et al., 2012) and categorized each according to the six basic emotions (and 'don't know'). Additionally, participants rated the intensity of each facial expression on a five point scale from 1 ('very weak') to 5 ('very strong'). Following the experiment, Jack et al. (2014) derived 720 individual facial expression models [60 participants $\times$ 6 emotions $\times$ 2 sex of face] using Pearson correlation. Specifically, for each AU, they correlated its binary activation vector (on/off) and participant responses to derive a 42 $\times$ 1 vector of AU activations. Next, they regressed each temporal parameter onto intensity ratings to obtain temporal parameter values for each significant AU. These experimental methods were therefore highly similar to those in chapter 4.

6.8 Supplementary tables

Table 6.1

Social trait shape PCA scores and correlations – Female

Social trait	**PC1**			**PC2**		
	mean (*SD*)	*r*	*p*	mean (*SD*)	*r*	*p*
Dom	-47.68 (30.13)	-.41	***	10.33 (17.25)	.13	0.29
Comp	-4.81 (29.35)	-.04	0.69	16.84 (12.51)	.21	*
Trustw	34.36 (23.95)	.28	***	21.47 (18.29)	.26	***
Warm	21.6 (22.14)	.18	*	38.8 (22.11)	.48	***
Subm	52.1 (36.64)	.45	***	-9.56 (19.43)	-.12	0.29
Incomp	7.59 (44.39)	.06	0.69	-21.71 (19.72)	-.27	*
Untrustw	-38.81 (26.67)	-.32	***	-19.95 (20.7)	-.24	***
Cold	-24.73 (25.55)	-.21	*	-36.31 (22.29)	-.45	***

Note. All *p*-values are based on Pearson's correlation and corrected for multiple comparisons with Bonferroni-Holm correction. * $p < .05$, *** $p < .001$

Table 6.2

Social trait shape PCA scores and correlations – Male

Social trait	PC1			PC2		
	mean (*SD*)	*r*	*p*	mean (*SD*)	*r*	*p*
Dom	-58.63 (24.36)	-.45	***	8.74 (17.56)	.11	0.3
Comp	-29.31 (27.72)	-.21	*	32.13 (22.02)	.39	***
Trustw	40.18 (24.55)	.29	***	22.13 (15.12)	.26	***
Warm	35.0 (21.01)	.27	***	31.77 (15.56)	.4	***
Subm	61.86 (30.6)	.47	***	-10.05 (18.68)	-.13	0.3
Incomp	30.79 (38.6)	.23	*	-37.24 (24.06)	-.45	***
Untrustw	-43.4 (20.46)	-.31	***	-20.85 (14.87)	-.25	***
Cold	-36.71 (20.07)	-.28	***	-26.84 (12.23)	-.34	***

Note. All *p*-values are based on Pearson's correlation and corrected for multiple comparisons with Bonferroni-Holm correction. * $p < .05$, *** $p < .001$

Table 6.3

Social trait complexion PCA scores and correlations – Female

Trait	PC1			PC2			PC3		
	mean (*SD*)	*r*	*p*	mean (*SD*)	*r*	*p*	mean (*SD*)	*r*	*p*
Dom	7.55 (8.56)	.25	***	-5.96 (6.37)	-.29	***	-3.33 (6.98)	-.18	.16
Comp	-2.75 (6.6)	-.09	1.96	-6.62 (4.75)	.43	***	-2.58 (3.99)	-.13	0.69
Trustw	-10.01 (7.3)	-.34	***	-1.84 (3.5)	-.09	1.96	-3.81 (5.73)	-.21	*
Warm	-9.04 (5.77)	-.31	***	-0.13 (4.39)	-.01	1.23	-6.5 (4.83)	-.36	***
Subm	-8.85 (8.75)	-.3	***	5.46 (6.85)	.27	***	3.37 (6.39)	.18	0.16
Incomp	3.92 (8.66)	.13	.78	9.19 (11.35)	.43	***	3.37 (5.9)	.17	0.19
Untrustw	10.28 (9.07)	.35	***	1.15 (4.5)	.06	3.1	3.29 (5.75)	.18	0.16
Cold	8.98 (7.79)	.31	***	-0.99 (4.31)	-.01	2.33	6.29 (4.36)	.35	***

Note. All *p*-values are based on Pearson's correlation and corrected for multiple comparisons with Bonferroni-Holm correction. * $p < .05$, *** $p < .001$

Table 6.4

Social trait complexion PCA scores and correlations – Male

Trait	PC1			PC2			PC3		
	mean (*SD*)	*r*	*p*	mean (*SD*)	*r*	*p*	mean (*SD*)	*r*	*p*
Dom	10.36 (7.53)	.34	***	-4.22 (5.67)	-.23	*	2.05 (4.46)	.14	0.58
Comp	2.40 (5.31)	.07	2.47	-5.81 (5.6)	-.29	***	0.79 (5.53)	.05	2.81
Trustw	-10.46 (8.45)	-.33	***	-3.55 (5.33)	-.18	0.15	-2.26 (7.45)	.13	0.52
Warm	-10.29 (8.87)	-.33	***	-4.55 (5.66)	-.24	*	-0.50 (5.42)	.05	1.43
Subm	-10.71 (8.84)	-.35	***	4.29 (5.90)	.23	*	-2.04 (4.95)	-.14	0.58
Incomp	-1.90 (6.14)	-.06	3.1	7.06 (6.26)	.35	***	-0.76 (7.37)	-.05	1.89
Untrustw	10.82 (7.97)	.34	***	2.81 (5.07)	.14	0.54	1.98 (4.47)	.13	0.73
Cold	9.86 (7.13)	.32	***	4.04 (5.63)	.22	*	0.73 (4.76)	.05	2.75

Note. All *p*-values are based on Pearson's correlation and corrected for multiple comparisons with Bonferroni-Holm correction. * $p < .05$, *** $p < .001$

6.9 Supplementary figures

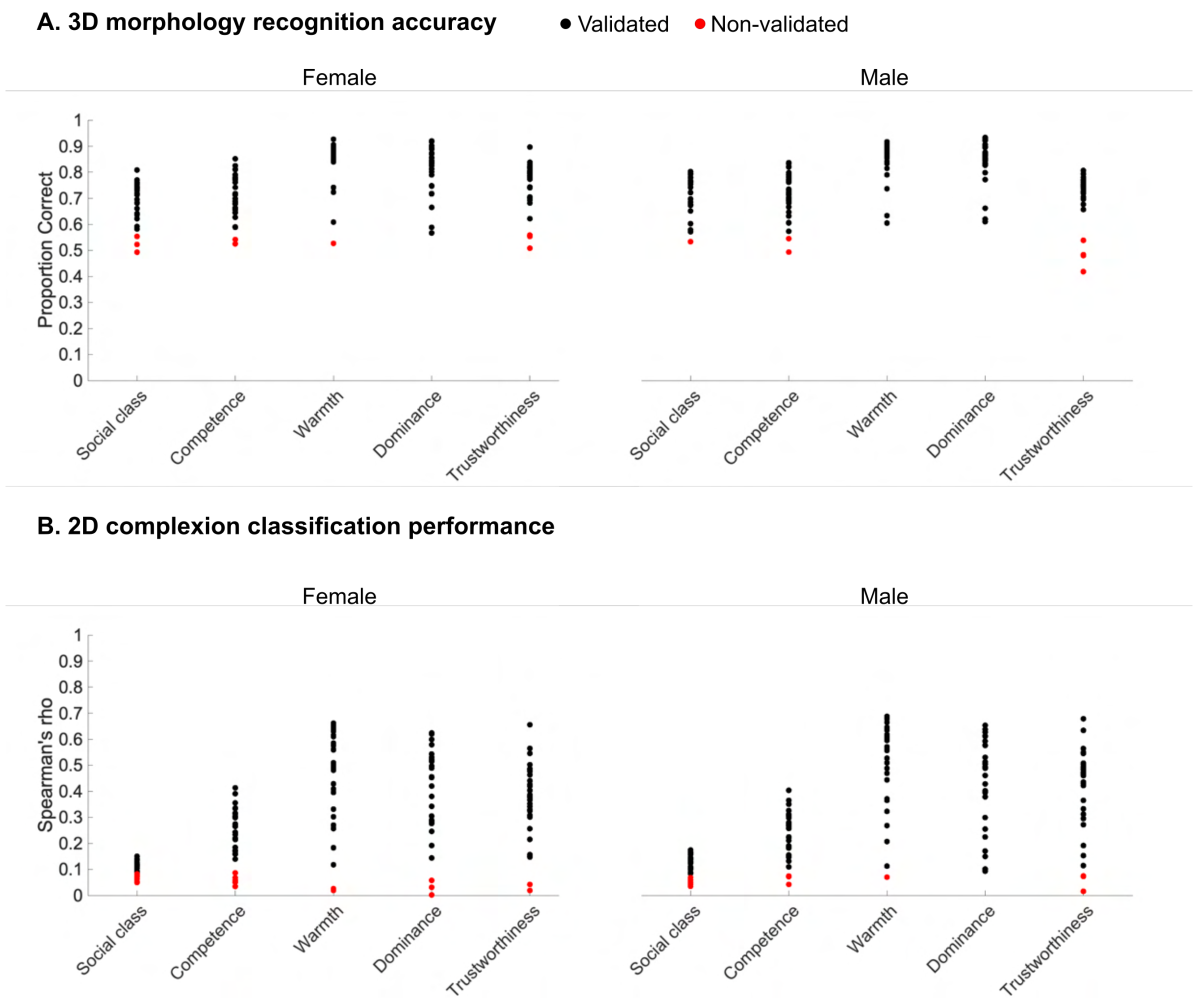

Figure 6.10. Validation of social class and social trait face models. Each subplot shows the performance of the social class and social trait face models for 3D shape and 2D complexion separately, and for each sex of stimulus face, tested under a validation procedure. Each point represents an individual participant face model. Black shows models that performed statistically significantly above chance, red shows models that did not. (A) 3D shape recognition accuracy. Each point shows the recognition accuracy of each model obtained in a 2-alternative-forced-choice validation task. Participants accurately recognized a total of 284 (95%) models (141 female: 27 social class, 28 competence, 29 warmth, 30 dominance, 27 trustworthiness; 143 male: 29 social class, 28 competence, 30 warmth, 30 dominance, 26 trustworthiness). (B) 2D complexion classification performance. Each point shows the classification accuracy, measured as Spearman's ρ, obtained in the leave-one-out cross-validation task. I obtained accurate classifications for 267 (89%) models (130 female: 22 social class, 25 competence, 28 warmth, 27 dominance, 28 trustworthiness; 137 male: 25 social class, 26 competence, 29 warmth, 30 dominance, 27 trustworthiness).

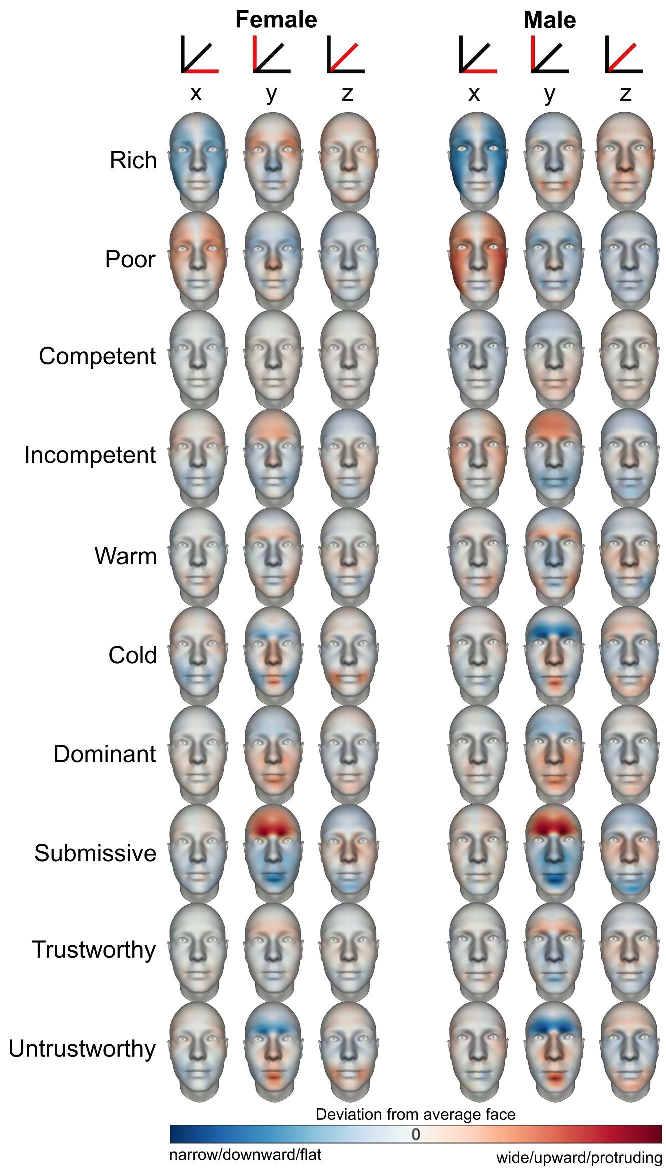

Figure 6.11. Average feature deviations of the social class and social trait face models from the average face. In each row, color-coded face maps show the feature deviations from the average face for a given social judgment, averaged across participants (validated models only) for each sex of stimulus faces and each x, y, and z plane separately. Red indicates positive deviations; blue indicates negative deviations (see colorbar at bottom). Specifically, in the x dimension, red represents outward (wider) deviations, blue represents inward (narrower) deviations. In the y dimension, red represents upward deviations, blue represents downward deviations. In the z dimension, red represents protruding deviations, blue represents flattening deviations. For example, faces rated as poor are wider (x) with shorter noses, lower eyes/brows, downturned mouths, shorter chins (y), and flatter (z) than the average face. Values are normalized across all face maps.

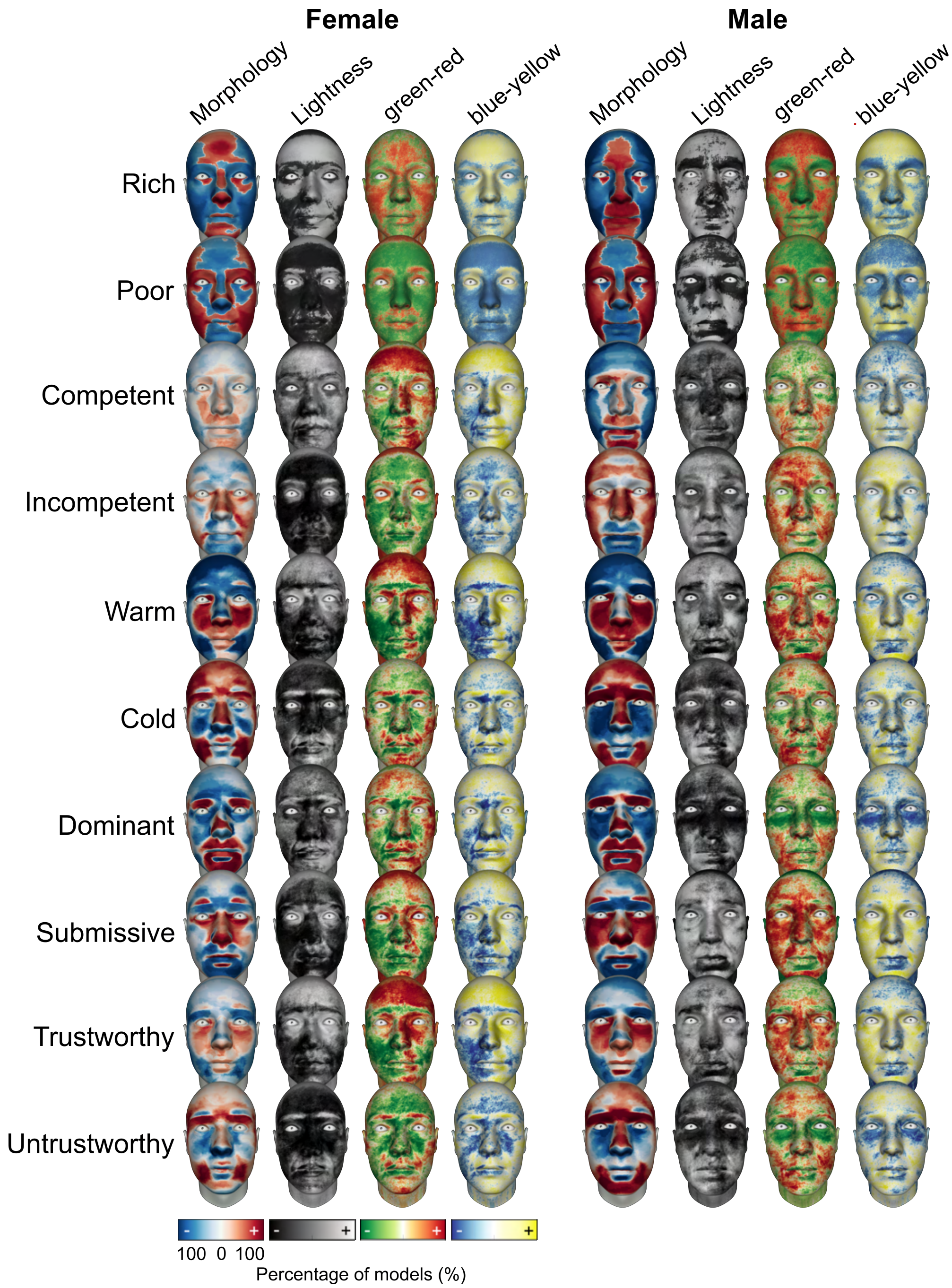

Figure 6.12. Replication of results across participants. In each row, color-coded faces show the percentage of participants (validated models only) whose face models comprise feature deviations from the average face in the same direction (positive or negative) for each vertex for 3D shape and for each pixel for 2D complexion (for each L*a*b color channel separately). For example, for rich male face models, most participants' 3D shape models comprised a narrow face with a more protruding nose and mouth. Similarly, for poor female face models most participants' 2D complexion models comprised darker, greener, and bluer (i.e., cooler) complexions.

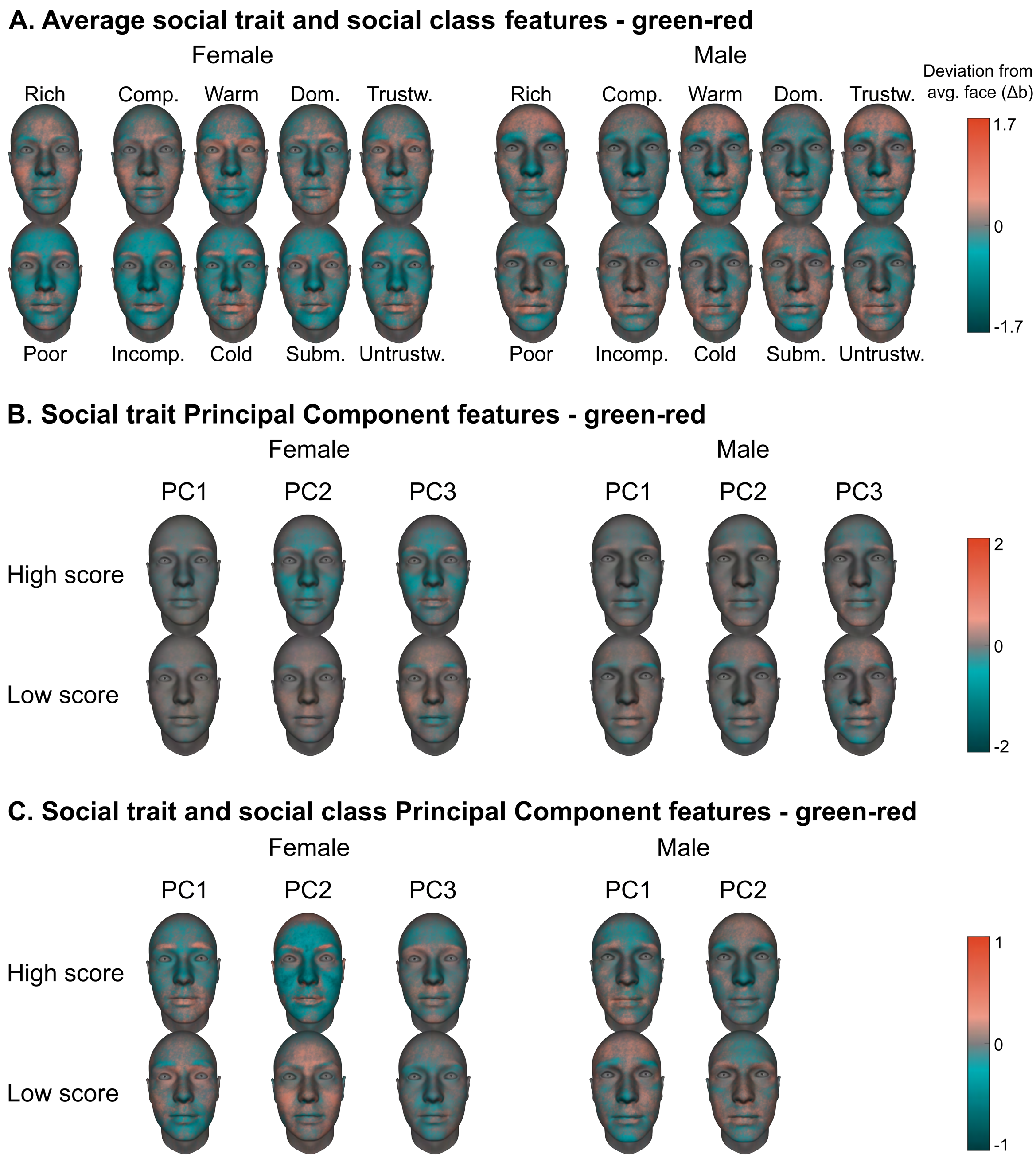

Figure 6.13. Main complexion results with colorblind friendly colors. (A) Average social trait and social class features - green-red. Face maps show complexion face features in the green-red color channel for each social class and social trait and each sex of stimulus face (normalized across all judgments). Red shows positive deviation from the average face; green shows negative deviation from the average face (see colorbar to right). (B) Social trait Principal Component features - green-red. Face maps show the complexion features captured by each social trait PC in the green-red color channel. Colormap as for A (see colorbar to right). (C) Social trait and social class Principal Component features - green-red. Face maps show the complexion features captured by each social class-social trait PC in the green-red color channel. Colormap as for A (see colorbar to right).

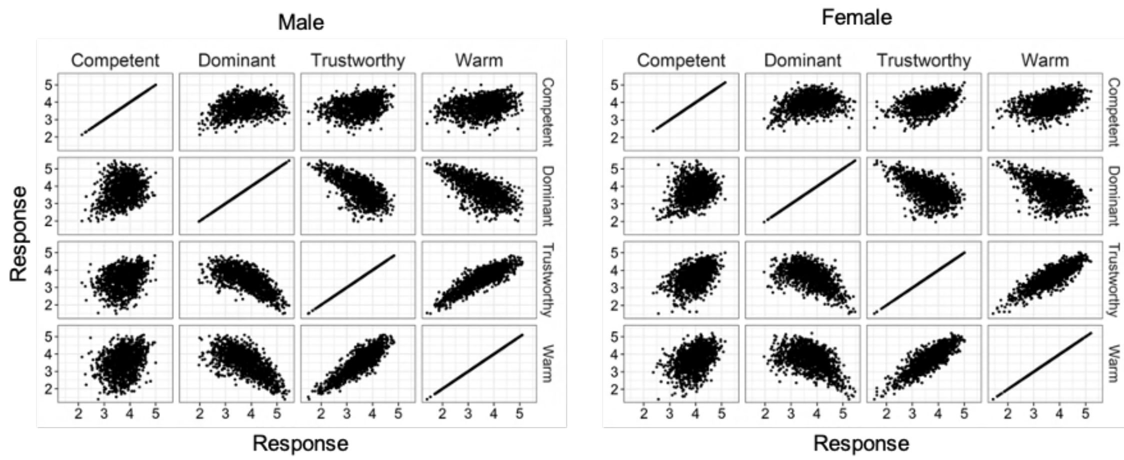

Figure 6.14. Social trait judgment correlations. Scatterplots show the correlation between mean behavioral ratings across participants of each social trait pair. Each black point represents one stimulus (1,200 stimuli per sex of face and social trait).

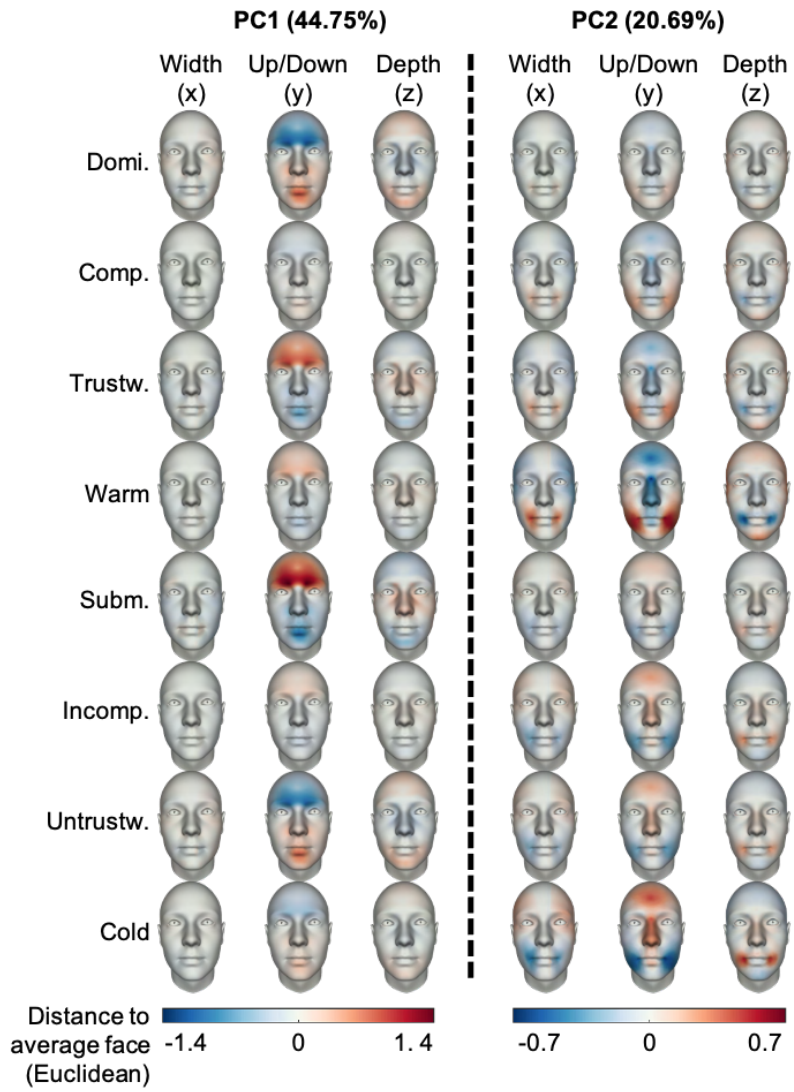

Figure 6.15. Shape principal component features – female. Colored face maps show the predicted shape effects for each one of the eight social traits, and each PC separately. Blue corresponds to narrow/downward/flat features (measured in Euclidean Distance to the average face). Red corresponds to wide/upward/protruding features. I also show the percentage of variance explained by each PC in parentheses.

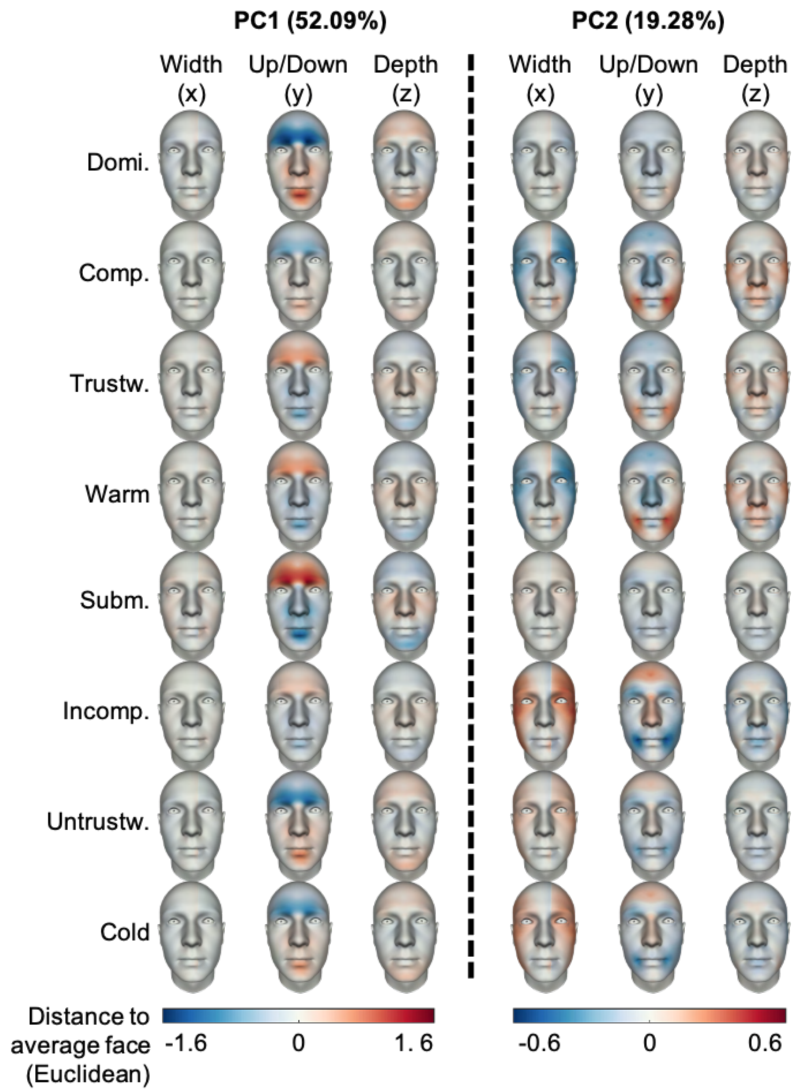

Figure 6.16. Shape principal component features – male. Colored face maps show the predicted shape effects for each one of the eight social traits, and each PC separately. Blue corresponds to narrow/downward/flat features (measured in Euclidean Distance to the average face). Red corresponds to wide/upward/protruding features. I also show the percentage of variance explained by each PC in parentheses.

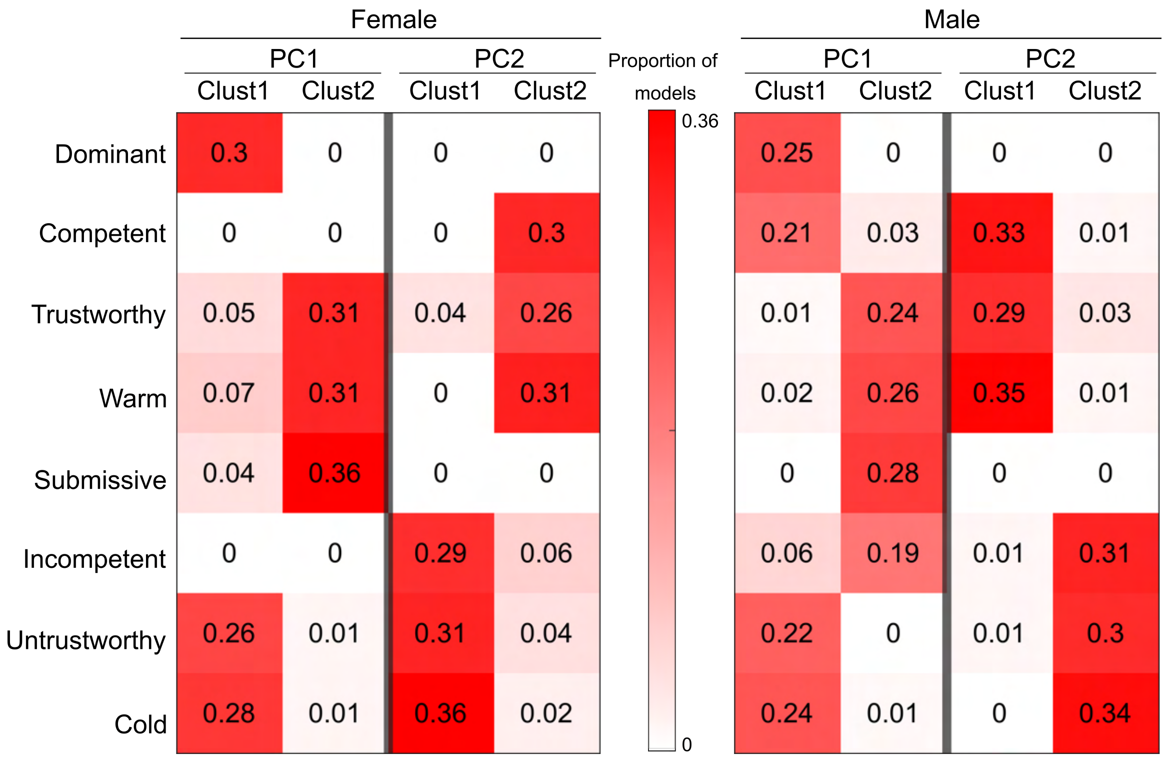

Figure 6.17. Social trait shape PCA clusters. Proportion of social trait models in each cluster (K-means of PCA scores). Red colors indicate higher proportion (see colorbar in center). For example, 30% of all social trait models in female PC1 cluster 1 (see labels at top) were dominant models.

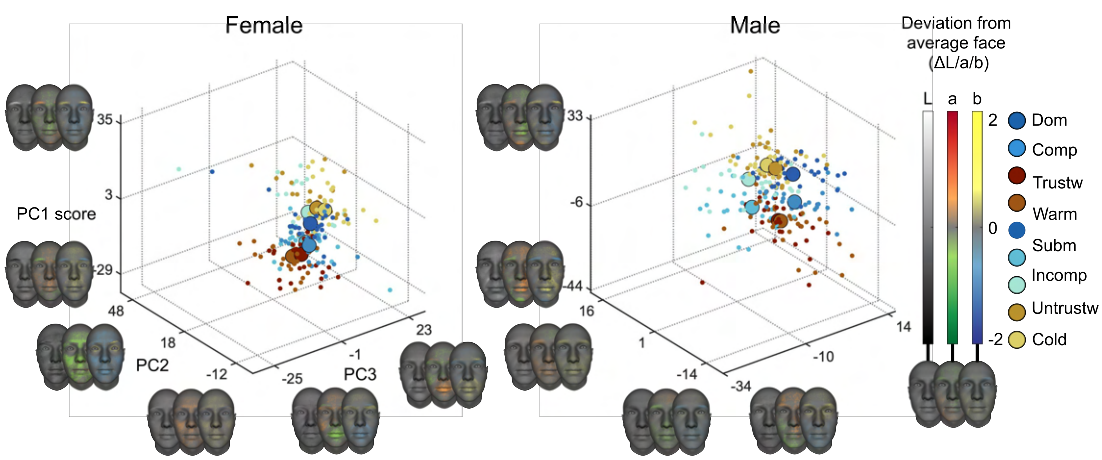

Figure 6.18. Social trait complexion latent feature space. Colored face maps show the complexion features of each PC compared to the average face for each color channel – Lightness (L), red-green (a), yellow-blue (b). Scatterplots show the distribution of individual participant face models (small points) according to their PC scores with larger point size indicating median scores across participants.

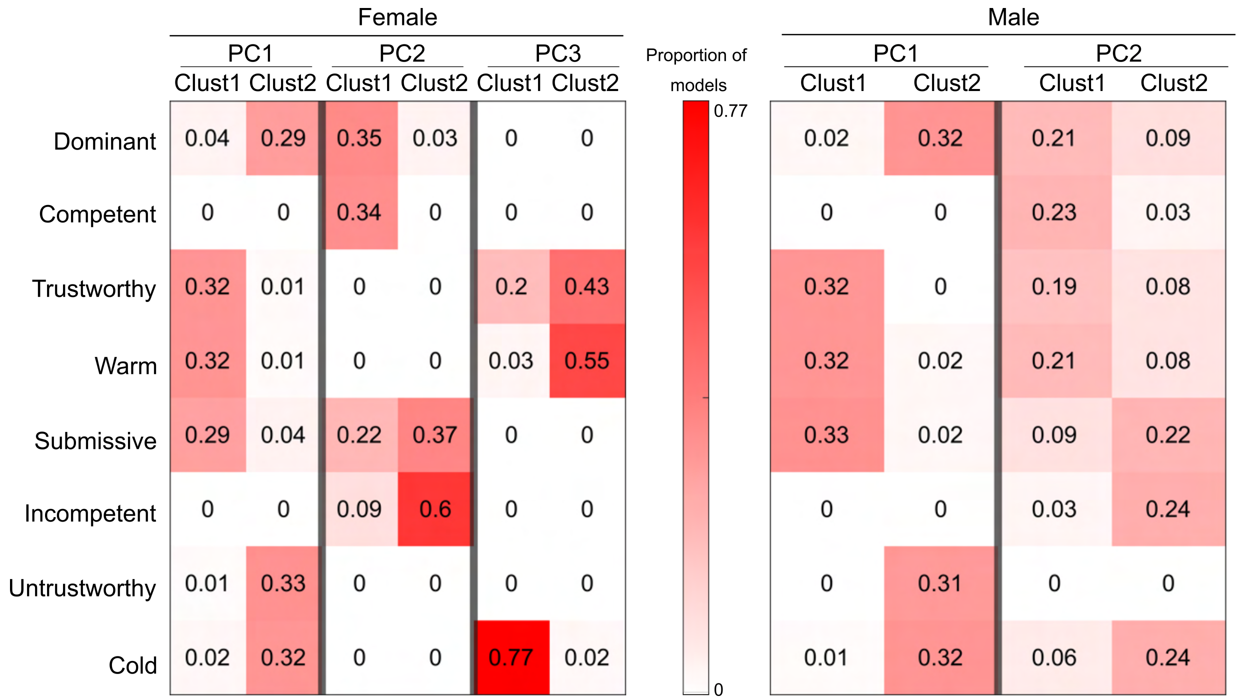

Figure 6.19. Social trait complexion PCA clusters. Proportion of social trait models in each PCA cluster (K-means of PCA scores). Red colors indicate higher proportion (see colorbar in center). For example, 32% of all social trait models in female PC1 cluster 1 (see labels at top) were trustworthy models.

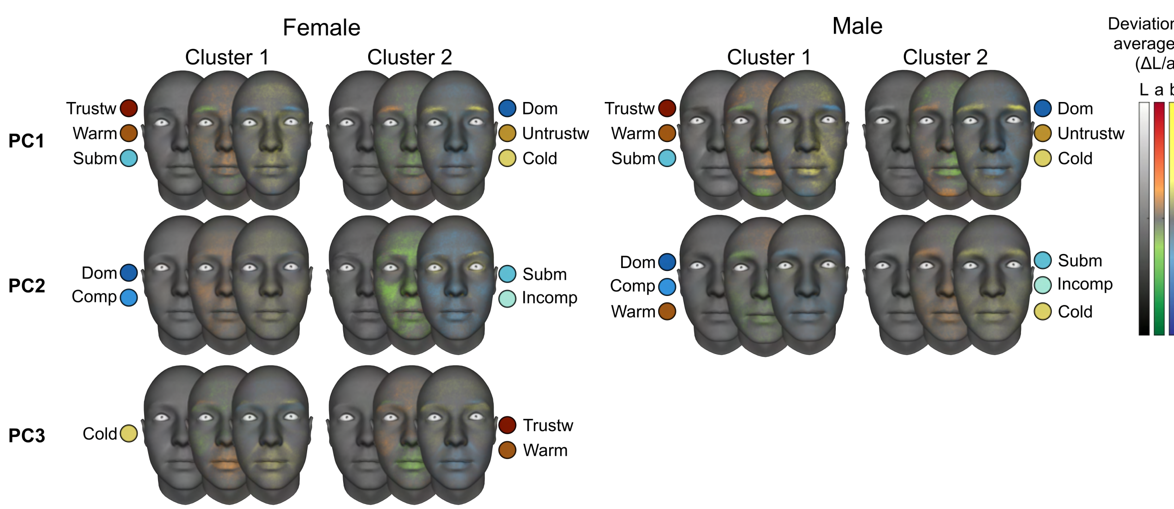

Figure 6.20. Social trait complexion PCA cluster features in CIELAB color space. Each row of faces shows the complexion features (in L*a*b; see colorbars to right) associated with each cluster (derived using K-means) of social traits (see labels) for one PC and for female and male faces separately. Each row corresponds to a different PC.

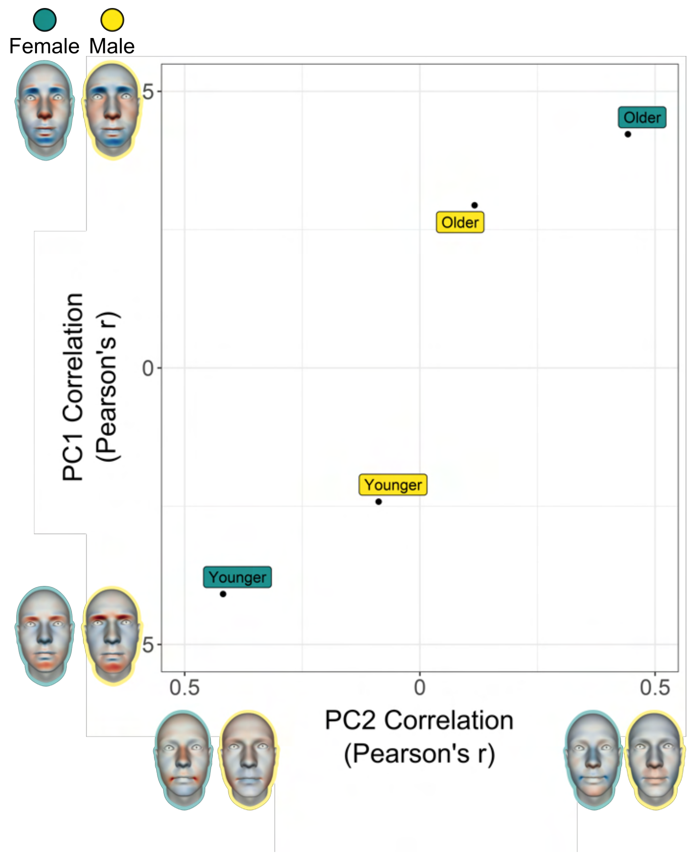

Figure 6.21. Correlations between latent social trait face shape features and age-related features. A scatterplot shows the correlation (Pearson's *r*) between each social trait shape PC (PC1 – y-axis; PC2 – x-axis) and younger and older adult face shape features. Colored points (female = green; male = yellow) show positive correlations between shape values of each social trait PC and younger (18 years) or older (35 years) faces (based on aging of the average female and male face shape). Colored face maps show the face features of each social trait PC. Negative correlations are excluded for clearer visualization.

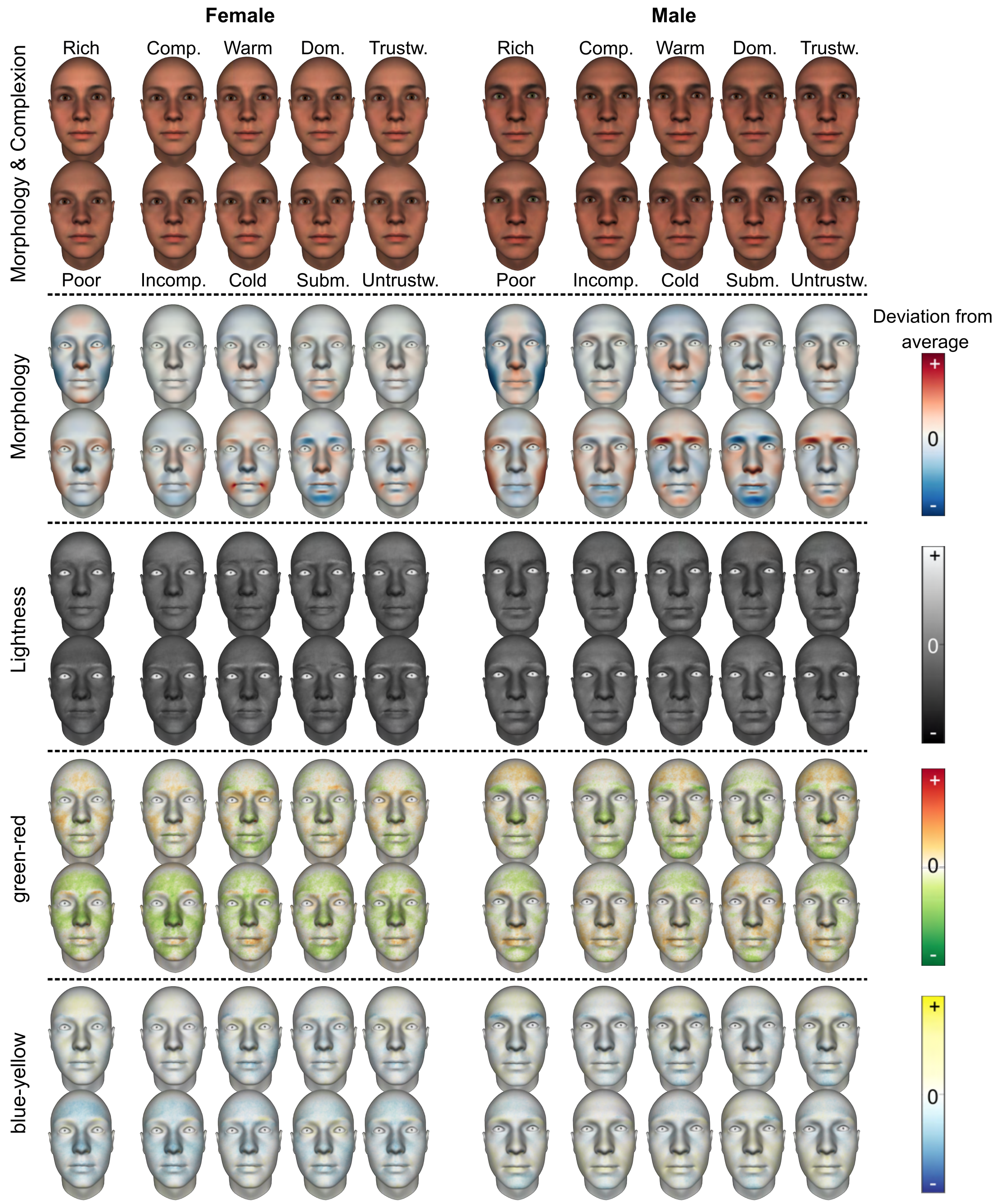

Figure 6.22. Average 3D face models of social class and social traits. Faces in the top row show the average predicted face models for social class each social trait and for each sex of stimulus face. Below, results are shown separately for 3D shape and 2D complexion. For shape, color-coding shows the direction of deviation from the average face where red represents positive deviation (e.g., poor faces are wider) and blue represents negative deviation (e.g., rich faces are narrower; see color bars on right). For complexion, color-coding represents the direction and magnitude of the deviation from the average face in L*a*b color space separately (e.g., rich faces have redder cheeks). Values are normalized across all social judgments and for shape and complexion separately for display purposes.

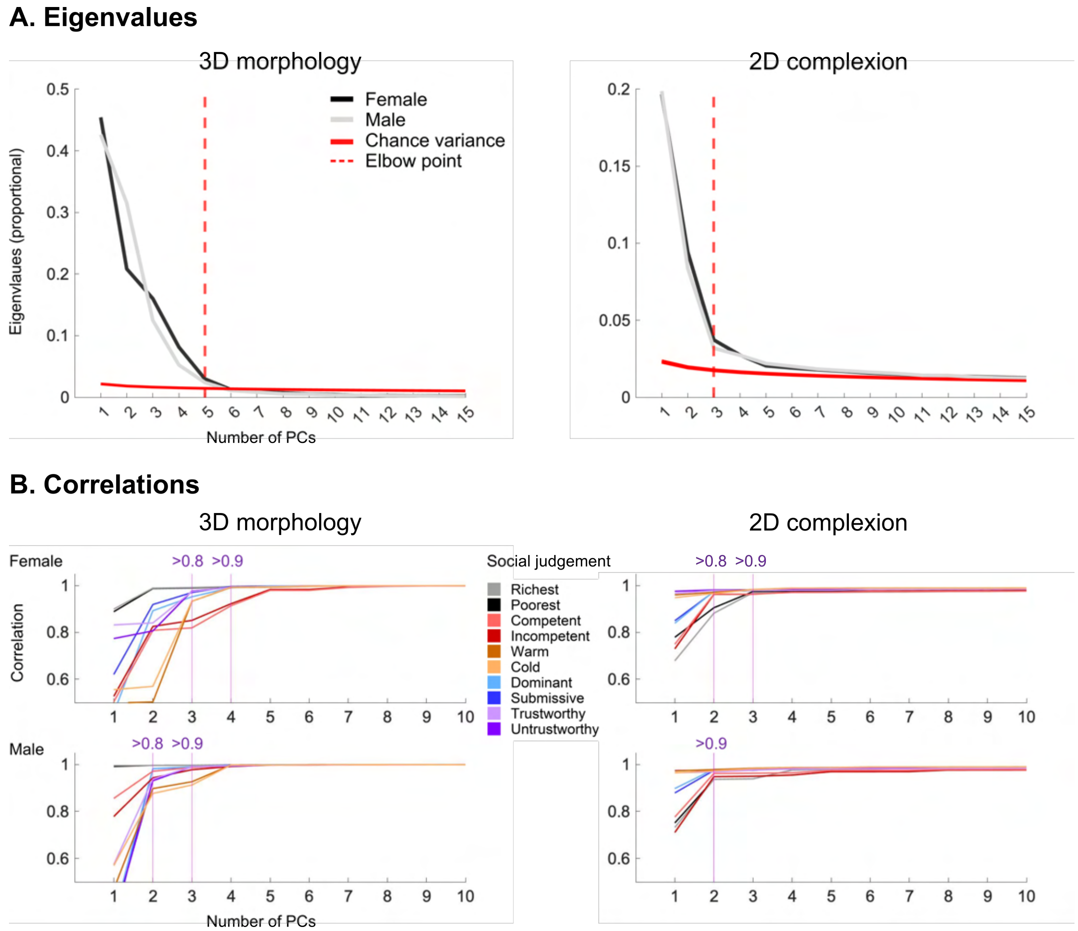

Figure 6.23. Social class Principal Component selection. Each panel shows the results of two complementary analyses used to estimate the optimal number of Principal Components (PCs) to represent all social class and social trait face models, for 3D shape and 2D complexion separately and for each sex of stimulus face. (A) Eigenvalues. In each plot, gray and black lines show the proportional eigenvalues (eigenvalue/sum of eigenvalues) across PCs for female (black) and male (gray) face models. The horizontal red line indicates the variance accounted for by each PC expected by chance where the broken stick method identifies the PCs that explain more variance than would be expected by chance (e.g., Cangelosi & Goriely, 2007; Caron, 2016). The vertical dashed red line indicates the elbow point (i.e., the point at which the curve of Eigenvalues visibly bends (e.g., Zambelli, 2016)). Results indicate 5 PCs for 3D shape and 3 for 2D complexion for male and female faces. (B) Correlations. Color-coded lines show the average correlation between the face models reconstructed based on n PCs (see x axis) vs all PCs for all social judgments (see legend). For 3D shape, results indicate that 4 PCs for female face models (85.18% variance accounted for) and 3 PCs for male face models (80.88% variance accounted for) are sufficient to obtain correlations above 0.9 (indicated by right-most magenta line) for all social judgments. For 2D complexion, results showed that 3 PCs for female face models (32.83% variance accounted for) and 2 PCs for male face models (28.25% variance accounted for) are sufficient to reach correlations above 0.9.

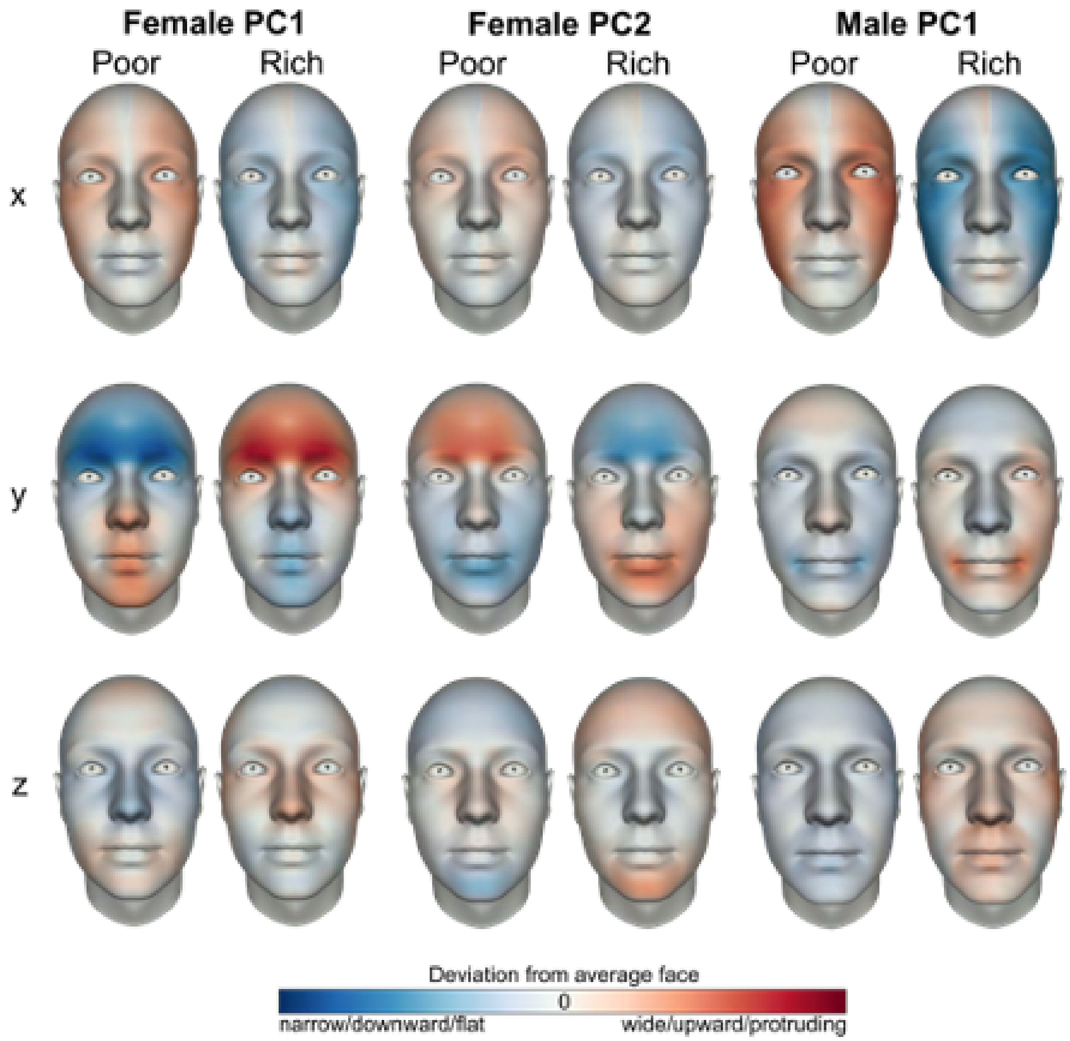

Figure 6.24. Shape Principal Components represented in x, y, and z dimensions separately, with social class as an example. Colored face maps show the shape features captured by each PC that is significantly associated with representing the social class face models, for female and male faces separately. Red indicates positive deviations from the average face; blue indicates negative deviations. In the x dimension, red represents outward deviations; blue represents inward deviations. In the y dimension, red represents upward deviations; blue represents downward deviations. In the z dimension, red represents protruding deviations; blue represents flattening deviations. For example, for PC1 for male faces models, poor faces are wider (x – red), have shorter noses, lower eyes/brows, downturned mouths, and shorter chins (y – blue), and are flatter (z – blue). Rich face models show the opposite pattern. Values are normalized across all face maps.

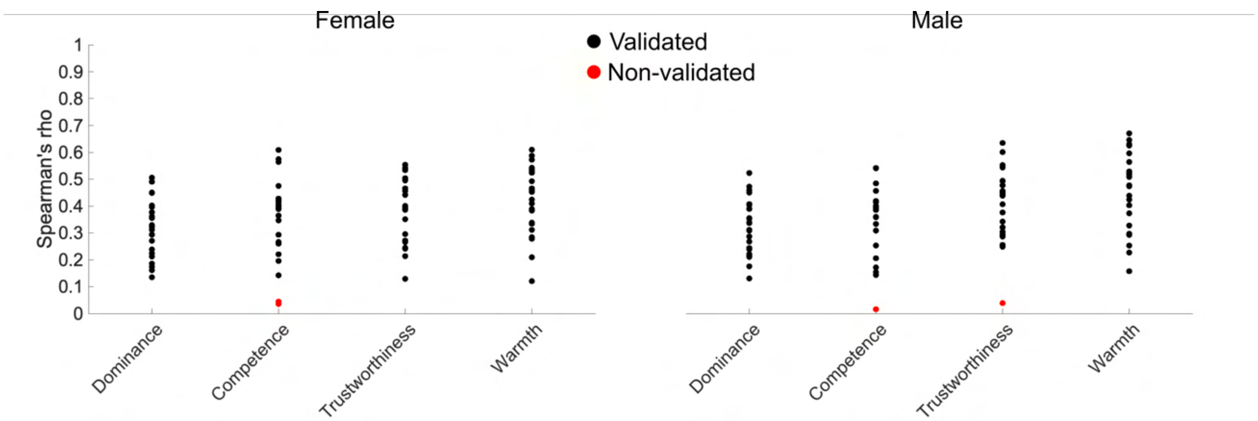

Figure 6.25. Validation of social trait facial expression models. Each subplot shows the validation performance of the social trait facial expression models for each sex of stimulus face. Each point shows the classification accuracy, measured as Spearman's ρ, obtained in the leave-one-out cross-validation task. I obtained accurate classifications for 178 (97.8%) models (89 female, 89 male; 46 dominance, 41 competence, 45 trustworthiness, 46 warmth).

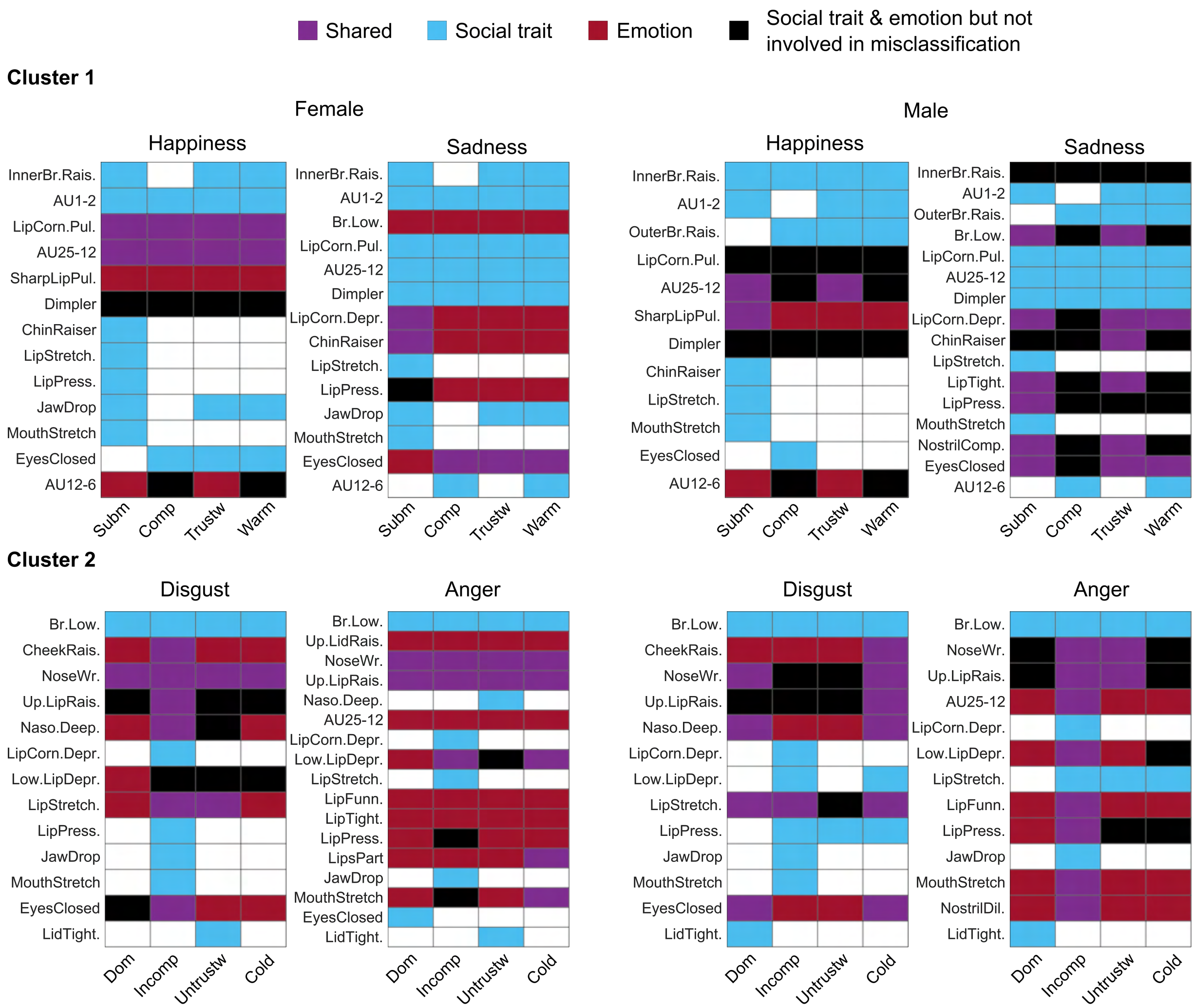

Figure 6.26. Social trait and emotional facial expression ablation results. Colored matrices show, for each sex of stimulus face, social trait cluster and each corresponding emotion the AUs that drove classifications of each social trait as the emotion in purple (based on ablation analysis). Cyan indicates AUs that were significant (> population prevalence) for the social trait only and red indicates AUs that were significant (> population prevalence) for the emotion only. AUs that were significant for the social trait and emotion but did not drive classification of one as the other are shown in black.

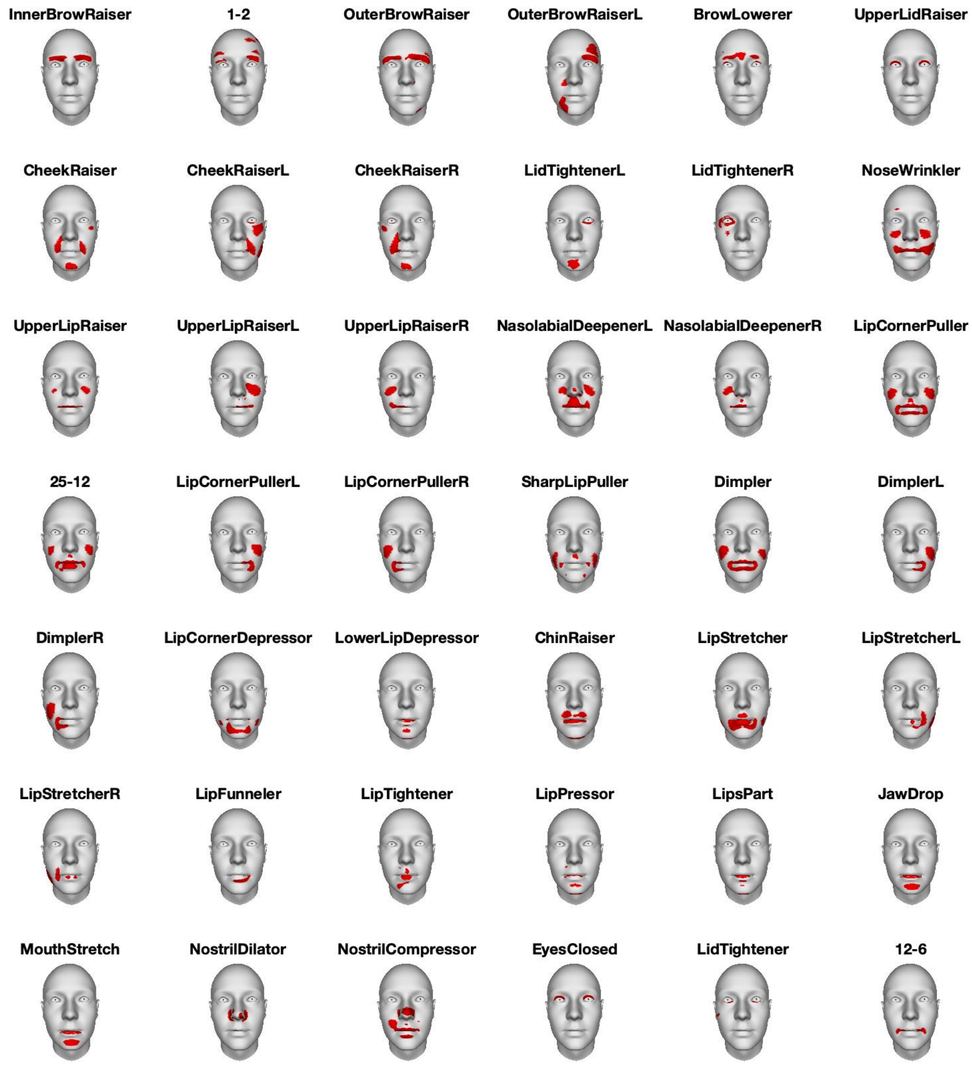

Figure 6.27. Face shape vertices of each Action Unit. For each AU (see titles in black), red areas on the face map show the 3D shape vertices that are most relevant to this AU (>85th percentile of distances to the average (Euclidean)).

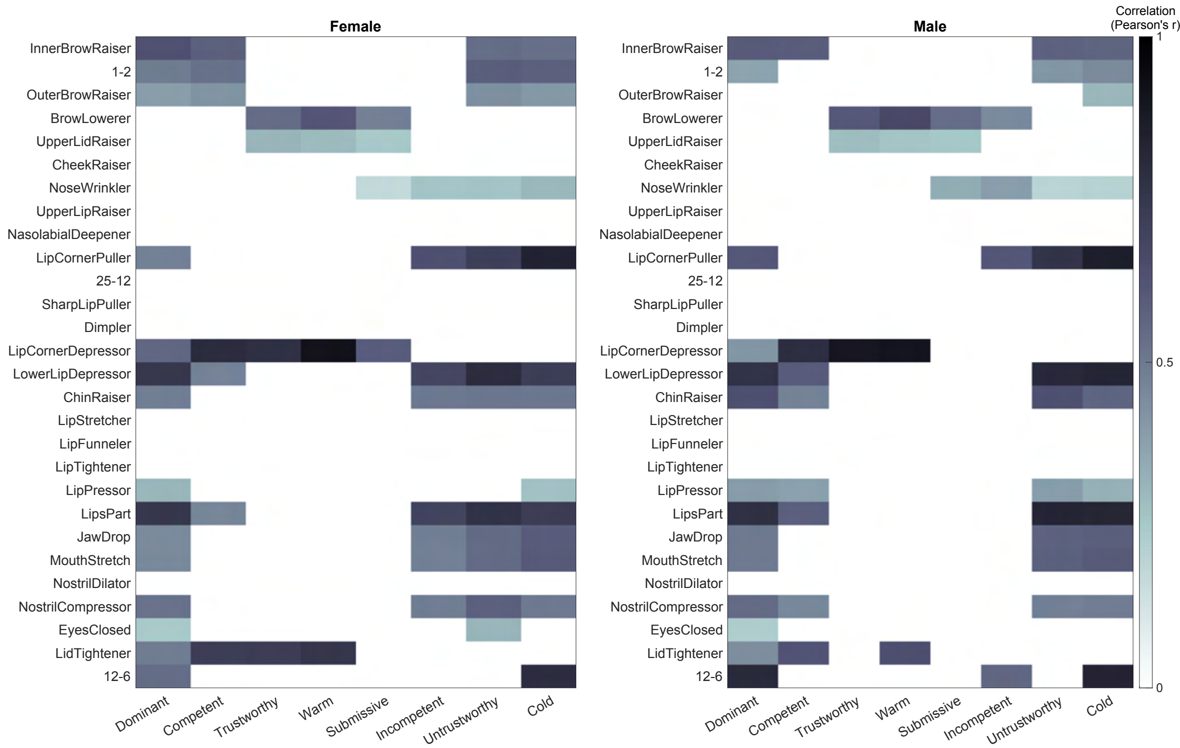

Figure 6.28. Correlations between individual Action Units and social trait face shapes. For each sex of stimulus face, gray-scale matrices show the correlation of 3D shape vertex residuals (from average) between each AU (y-axis) and social trait shape (x-axis; mean across individual participant comparisons). Only significant positive correlations ($p \leq .05$; corrected with Bonferroni-Holm correction) that also passed the population prevalence threshold (Ince et al., 2021) are displayed (others set to zero). Darker colors correspond to stronger correlations. For example, Inner Brow Raiser positively correlated with dominant for female faces. See also Figure 4.4 for further details on how these correlation values were derived.

References

Abele, A. E., Amy Cuddy, J. C., Judd, C. M., & Yzerbyt, V. Y. (2008). Fundamental dimensions of social judgment. *European Journal of Social Psychology*, *38*(7), 1063–1065. https://doi.org/10.1002/EJSP.574

Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness—communion with warmth and morality. *Frontiers in Psychology*, 7:1810. https://doi.org/10.3389/FPSYG.2016.01810

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of personality and social psychology*, *93*(5), 751–763. https://doi.org/10.1037/0022-3514.93.5.751

Adams, R. B., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion*, *5*(1), 3–11. https://doi.org/10.1037/1528-3542.5.1.3
237

Adams, R. B., Albohn, D. N., Kveraga, K., Reginald B. Adams, J., Albohn, D. N., & Kveraga, K. (2017). Social vision: Applying a Social-Functional approach to face and expression perception: *https://doi.org/10.1177/0963721417706392*, *26*(3), 243–248. https://doi.org/10.1177/0963721417706392

Adams, R. B., Hess, U., & Kleck, R. E. (2015). The intersection of gender-related facial appearance and facial displays of emotion. *Emotion Review*, *7*(1), 5–13. https://doi.org/10.1177/1754073914544407

Adams, R. B., Nelson, A. J., Soto, J. A., Hess, U., & Kleck, R. E. (2012). Emotion in the neutral face: A mechanism for impression formation? *Cognition and Emotion*, *26*(3), 431–441. https://doi.org/10.1080/02699931.2012.666502

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Syme, S. L. (1994). Socioeconomic status and health: The challenge of the gradient. *American Psychologist*, *49*(1), 15–24. https://doi.org/10.1037/0003-066X.49.1.15

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning:

Preliminary data in healthy white women. *Health Psychology*, *19*(6), 586–592. https://doi.org/10.1037/0278-6133.19.6.586

Aguinaldo, E. R., & Peissig, J. J. (2021). Who's behind the makeup? The effects of varying levels of cosmetics application on perceptions of facial attractiveness, competence, and sociosexuality. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.661006

Ahler, D. J., Citrin, J., Dougal, M. C., & Lenz, G. S. (2016). Face value? Experimental evidence that candidate appearance influences electoral choice. *Political Behavior 2016 39:1*, *39*(1), 77–102. https://doi.org/10.1007/S11109-016-9348-6

Ahumada, A. J., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, *49*, 1751–1756.

Al Moubayed, N., Vazquez-Alvarez, Y., McKay, A., & Vinciarelli, A. (2014). Face-based automatic personality perception. *MM 2014 - Proceedings of the 2014 ACM Conference on Multimedia*, 1153–1156. https://doi.org/10.1145/2647868.2655014

Albert, G., Wells, E., Arnocky, S., Liu, C. H., & Hodges-Simeon, C. R. (2021). Observers use facial masculinity to make physical dominance assessments following 100-ms exposure. *Aggressive Behavior*, *47*(2), 226–235. https://doi.org/10.1002/ab.21941

Albohn, D. N., & Adams, R. B. (2021). The Expressive Triad: Structure, Color, and Texture Similarity of Emotion Expressions Predict Impressions of Neutral Faces. *Frontiers in Psychology*, *12*, 220. https://doi.org/10.3389/fpsyg.2021.612923

Alkozei, A., Haack, M., Skalamera, J., Smith, R., Satterfield, B. C., Raikes, A. C., & Killgore, W. D. (2018). Chronic sleep restriction affects the association between implicit bias and explicit social decision making. *Sleep Health*, *4*(5), 456–462. https://doi.org/10.1016/j.sleh.2018.07.003

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. https://doi.org/10.1037/h0093360

Ambadar, Z., Cohn, J. F., & Reed, L. I. (2009). All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *Journal of Nonverbal Behavior*, *33*(1), 17–34. https://doi.org/10.1007/s10919-008-0059-5

Ames, D. R., Kammrath, L. K., Suppes, A., & Bolger, N. (2010). Not so fast: The (not-quite-complete) dissociation between accuracy and confidence in thin-slice impressions. *Personality and Social Psychology Bulletin*, *36*(2), 264–277. https://doi.org/10.1177/0146167209354519

Antonakis, J., & Eubanks, D. L. (2017). Looking leadership in the face. *Current Directions in Psychological Science*, *26*(3), 270–275. https://doi.org/10.1177/0963721417705888

Arnocky, S., Carré, J. M., Bird, B. M., Moreau, B. J., Vaillancourt, T., Ortiz, T., & Marley, N. (2018). The facial Width-to-Height ratio predicts sex drive, sociosexuality, and intended infidelity. *Archives of Sexual Behavior*, *47*(5), 1375–1385. https://doi.org/10.1007/s10508-017-1070-x

Asch, S. E. (1946). Forming Impressions of Personality. *The Journal of Abnormal and Social Psychology*, *41*(3), 258–290. https://doi.org/https://doi.org/10.1037/h0055756

Ashmore, R. D., & Tumia, M. L. (1980). Sex stereotypes and implicit personality theory. I. A personality description approach to the assessment of sex stereotypes. *Sex Roles*, *6*(4), 501–518. https://doi.org/10.1007/BF00287882

Aylett, R., Paiva, A., Dias, J., Hall, L., & Woods, S. (2009). Affective agents for education against bullying. *Affective information processing* (pp. 75–90). Springer, London. https://doi.org/10.1007/978-1-84800-306-4_5

Aznar-Casanova, J., Torro-Alves, N., & Fukusima, S. (2010). How much older do you get when a wrinkle appears on your face? Modifying age estimates by number of wrinkles. *Aging, Neuropsychology, and Cognition*, *17*(4), 406–421. https://doi.org/10.1080/13825580903420153

Bacev-Giles, C., & Haji, R. (2017). Online first impressions: Person perception in social media profiles. *Computers in Human Behavior*, *75*, 50–57. https://doi.org/10.1016/J.CHB.2017.04.056

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science*, *21*(3), 372–374. https://doi.org/10.1177/0956797609360756

Bailenson, J. N., Iyengar, S., Yee, N., & Collins, N. A. (2008). Facial similarity between voters and dandidates causes influence. *Public Opinion Quarterly*, *72*(5), 935–961. https://doi.org/10.1093/POQ/NFN064

Bakmazian, A. (2014). The man behind the beard: Perception of men's trustworthiness as a function of facial hair. *Psychology*, *2014*(03), 185–191. https://doi.org/10.4236/PSYCH.2014.53029

Balas, B., & Pacella, J. (2017). Trustworthiness perception is disrupted in artificial faces. *Computers in Human Behavior*, *77*, 240–248. https://doi.org/10.1016/j.chb.2017.08.045

Barber, S. J., Lee, H., Becerra, J., & Tate, C. C. (2019). Emotional expressions affect perceptions of younger and older adults' everyday competence. *Psychology and Aging*, *34*(7), 991–1004. https://doi.org/10.1037/pag0000405

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Bee, N., Pollock, C., André, E., & Walker, M. (2010). Bossy or wimpy: Expressing social dominance by combining gaze and linguistic behaviors. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes* in Bioinformatics), *6356 LNAI*, 265–271. https://doi.org/10.1007/978-3-642-15892-6_28

Bekhouche, S. E., Dornaika, F., Ouafi, A., & Taleb-Ahmed, A. (2017). Personality traits and job candidate screening via analyzing facial videos. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2017-July*, 1660–1663. https://doi.org/10.1109/CVPRW.2017.211

Berry, D. S., & McArthur, L. Z. (1985). Some components and consequences of a babyface. *Journal of Personality and Social Psychology*, *48*(2), 312–323. https://doi.org/10.1037/0022-3514.48.2.312

Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces. The impact of age-related craniofacial changes on social perception. *Psychological Bulletin*, *100*(1), 3–18. https://doi.org/10.1037/0033-2909.100.1.3

Biancardi, B., Cafaro, A., & Pelachaud, C. (2017). Could a virtual agent bewarm and competent? Investigating user's impressions of agent's non-verbal behaviours. *ISIAA 2017 - Proceedings of the 1st ACM SIGCHI International Workshop on Investigating Social Interactions with Artificial Agents, Co-located with ICMI 2017*, 22–24. https://doi.org/10.1145/3139491.3139498

Bin Meshar, M., Stolier, R. M., & Freeman, J. B. (2021). Facial Stereotyping Drives Judgments of Perceptually Ambiguous Social Groups. *Social Psychological and Personality Science*. https://doi.org/10.1177/19485506211062285

Bjornsdottir, R. T., & Rule, N. O. (2017). The visibility of social class from facial cues. *Journal of Personality and Social Psychology*, *113*(4), 530–546. https://doi.org/10.1037/PSPA0000091

Bjornsdottir, R. T., & Rule, N. O. (2020). Negative emotion and perceived social class. *Emotion (Washington, D.C.)*, *20*(6), 1031–1041. https://doi.org/10.1037/EMO0000613

Bjornsdottir, R. T. (2019). *Investigating the pervasiveness of social class cues in the face* (Doctoral dissertation). https://tspace.library.utoronto.ca/handle/1807/97322

Blair, I. V., Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and afrocentric facial features in social judgments. *Journal of Personality and Social Psychology*, *87*(6), 763–778. https://doi.org/10.1037/0022-3514.87.6.763

Bøggild, T., & Laustsen, L. (2016). An intra-group perspective on leader preferences: Different risks of exploitation shape preferences for leader facial dominance. *The Leadership Quarterly*, *27*(6), 820–837. https://doi.org/10.1016/J.LEAQUA.2016.09.003

Bolwig, N. (1978). Communicative signals and social behaviour of some African monkeys: A comparative study. *Primates*, *19*(1), 61–99. https://doi.org/10.1007/BF02373227

Borkenau, P., Brecke, S., Möttig, C., & Paelecke, M. (2009). Extraversion is accurately perceived after a 50-ms exposure to a face. *Journal of Research in Personality*, *43*(4), 703–706. https://doi.org/10.1016/J.JRP.2009.03.007

Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*(1), 55–59. https://doi.org/10.1016/j.anbehav.2011.03.024

Boshyan, J., Zebrowitz, L. A., Franklin, R. G., McCormick, C. M., & Carré, J. M. (2014). Age similarities in recognizing threat from faces and diagnostic cues. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *69*(5), 710–718. https://doi.org/10.1093/geronb/gbt054

Bradbury, J. W., & Vehrencamp, S. L. (2011). *Principles of animal communication, 2nd ed.* https://psycnet.apa.org/record/2011-20344-000

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brambilla, M., Biella, M., & Freeman, J. B. (2018). The influence of visual context on the evaluation of facial trustworthiness. *Journal of Experimental Social Psychology*, *78*, 34–42. https://doi.org/10.1016/j.jesp.2018.04.011

Brambilla, M., & Leach, C. W. (2014). On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition*, *32*(4), 397–408. https://doi.org/10.1521/SOCO.2014.32.4.397

Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology*, *41*(2), 135–143. https://doi.org/10.1002/EJSP.744

Brander, M., Egger, S. T., Hürlimann, N., Seifritz, E., Sumner, R. W., Vetter, S., & Magnenat, S. (2021). Virtual reality human-human interface to deliver psychotherapy to people experiencing auditory verbal hallucinations: Development and usability study. *JMIR Serious Games*, *9*(2), e26820. https://doi.org/10.2196/26820

Brescoll, V. L. (2016). Leading with their hearts? How gender stereotypes of emotion lead to biased evaluations of female leaders. *The Leadership Quarterly*, *27*(3), 415–428. https://doi.org/10.1016/J.LEAQUA.2016.02.005

Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualising mental representations: A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology*, *28*(1), 333–361. https://doi.org/10.1080/10463283.2017.1381469

Broadbent, E., Kumar, V., Li, X., Sollers III, J., Stafford, R. Q., MacDonald, B. A., & Wegner, D. M. (2013). Robots with display screens: A robot with a more humanlike face display is perceived to have more mind and a better personality. *Plos One*, *8*(8). https://doi.org/10.1371/journal.pone.0072589
20

Brown-Iannuzzi, J. L., Dotsch, R., Cooley, E., & Payne, B. K. (2017). The relationship between mental representations of welfare recipients and attitudes toward welfare. *Psychological Science*, *28*(1), 92–103. https://doi.org/10.1177/0956797616674999

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* University of California Press. https://doi.org/10.2307/2216617

Buchan, N. R., Croson, R. T., & Solnick, S. (2008). Trust and gender: An examination of behavior and beliefs in the Investment Game. *Journal of Economic Behavior & Organization*, *68*(3-4), 466–476. https://ideas.repec.org/a/eee/jeborg/v68y2008i3-4p466-476.html

Burgoon, J. K., Birk, T., & Pfau, M. (1990). Nonverbal behaviors, persuasion, and credibility. *Human Communication Research*, *17*(1), 140–169. https://doi.org/10.1111/j.1468-2958.1990.tb00229.x

Burriss, R. P., Little, A. C., & Nelson, E. C. (2007). 2D:4D and sexually dimorphic facial characteristics. *Archives of Sexual Behavior*, *36*(3), 377–384. https://doi.org/10.1007/s10508-006-9136-1

Burt, A. L., & Crewther, D. P. (2020). The 4D space-time dimensions of facial perception. *Frontiers in psychology*, *11*. https://doi.org/10.3389/FPSYG.2020.01842

Calvo, M. G., Gutiérrez-García, A., Avero, P., & Lundqvist, D. (2013). Attentional mechanisms in judging genuine and Fake smiles: Eye-movement patterns. *Emotion*, *13*(4), 792–802. https://doi.org/10.1037/A0032317

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, *11*(12), 535–543. https://doi.org/10.1016/J.TICS.2007.10.001

Cangelosi, R., & Goriely, A. (2007). Component retention in principal component analysis with application to cDNA microarray data. *Biology Direct*, *2*. https://doi.org/10.1186/1745-6150-2-2

Caron, P. O. (2016). A Monte Carlo examination of the broken-stick distribution to identify components to retain in principal component analysis. *Journal of Statistical Computation and Simulation*, *86*(12), 2405–2410. https://doi.org/10.1080/00949655.2015.1112390

Carragher, D. J., Thomas, N. A., & Nicholls, M. E. (2021). The dissociable influence of social context on judgements of facial attractiveness and trustworthiness. *British Journal of Psychology*, *112*(4), 902–933. https://doi.org/10.1111/bjop.12501

Carre, J. M., Campbell, J. A., Lozoya, E., Goetz, S. M. M., & Welker, K. M. (2013). Changes in testosterone mediate the effect of winning on subsequent aggressive behaviour. *Psychoneuroendocrinology*, *38*(10), 2034–2041. https://doi.org/10.1016/j.psyneuen.2013.03.008

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological science*, *20*(10), 1194–1198. https://doi.org/10.1111/J.1467-9280.2009.02423.X

Carré, J. M., Morrissey, M. D., Mondloch, C. J., & McCormick, C. M. (2010). Estimating aggression from emotionally neutral faces: Which facial cues are diagnostic? *Perception*, *39*(3), 356–377. https://doi.org/10.1068/p6543

Carrier, A., Louvet, E., Chauvin, B., & Rohmer, O. (2014). The primacy of agency over competence in status perception. *Social Psychology*, *45*(5), 347–356. https://doi.org/10.1027/1864-9335/a000176

Carrito, M. L., Santos, I. M., Bem-Haja, P., Lopes, A. A., Silva, C. F., & Perrett, D. I. (2020). The attractive side of trustworthiness: Effects of relationship context and social interaction anxiety on face preferences. *Evolutionary Behavioral Sciences*, *14*(3), 261–269. https://doi.org/10.1037/ebs0000177

Carroll, J. S., & Anderson, N. H. (1982). Foundations of Information Integration Theory. *The American Journal of Psychology*, *95*(4), 708. https://doi.org/10.2307/1422202

Chan-Serafin, S., Wang, L., & Wang, L. (2019). The role of positive emotion expression on perceived warmth and competence of women and men at work. *Academy of Management Proceedings*, *2019*(1), 16227. https://doi.org/10.5465/ambpp.2019.16227abstract

Charbonneau, I., Robinson, K., Blais, C., & Fiset, D. (2020). Implicit race attitudes modulate visual information extraction for trustworthiness judgments. *PLoS ONE*, *15*(9 September). https://doi.org/10.1371/journal.pone.0239305

Chaudhuri, A., & Gangadharan, L. (2007). An experimental analysis of trust and trustworthiness. *Southern Economic Journal*, *73*(4), 959–985. https://doi.org/10.1002/j.2325-8012.2007.tb00813.x

Chen, F. F., Jing, Y., & Lee, J. M. (2014). The looks of a leader: Competent and trustworthy, but not dominant. *Journal of Experimental Social Psychology*, *51*, 27–33. https://doi.org/10.1016/J.JESP.2013.10.008

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

Chiao, J. Y. (2010). Neural basis of social status hierarchy across species. *Current Opinion in Neurobiology*, *20*(6), 803–809. https://doi.org/10.1016/j.conb.2010.08.006

Chua, K. W., & Freeman, J. B. (2022). Learning to judge a book by its cover: Rapid acquisition of facial stereotypes. *Journal of Experimental Social Psychology*, *98*, 104225. https://doi.org/10.1016/j.jesp.2021.104225

Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from faces: A developmental study. *Psychological Science*, *25*(5), 1132–1139. https://doi.org/10.1177/0956797614523297

Cone, J., Brown-Iannuzzi, J. L., Lei, R., & Dotsch, R. (2020). Type I error is inflated in the two-phase reverse correlation procedure. *Social Psychological and Personality Science*, 1–9. https://doi.org/10.1177/1948550620938616

Connolly, H. L., Young, A. W., & Lewis, G. J. (2021). Face perception across the adult lifespan: evidence for age-related changes independent of general intelligence. *Cognition and Emotion*, *35*(5), 890–901. https://doi.org/10.1080/02699931.2021.1901657

Connor, P., Varney, J., Keltner, D., & Chen, S. (2021). Social class competence stereotypes are amplified by socially signaled economic inequality. *Personality and Social Psychology Bulletin*, *47*(1), 89–105. https://doi.org/10.1177/0146167220916640

Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The Stereotype Content Model and the BIAS map. *Advances in Experimental Social Psychology*, *40*, 61–149. https://doi.org/10.1016/S0065-2601(07)00002-0

Cuddy, A. J., Glick, P., & Beninger, A. (2011). The dynamics of warmth and competence judgments, and their outcomes in organizations. https://doi.org/10.1016/j.riob.2011.10.004

Cuddy, A. J., Norton, M. I., & Fiske, S. T. (2005). This old stereotype: The pervasiveness and persistence of the elderly stereotype. *Journal of Social Issues*, *61*(2), 267–285. https://doi.org/10.1111/J.1540-4560.2005.00405.X

Curtis, V., de Barra, M., & Aunger, R. (2011). Disgust as an adaptive system for disease avoidance behaviour. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *366*(1563), 389–401. https://doi.org/10.1098/rstb.2010.0117

De Neys, W., Hopfensitz, A., & Bonnefon, J. F. (2017). Split-second trustworthiness detection from faces in an economic game. *Experimental Psychology*, *64*(4), 231–239. https://doi.org/10.1027/1618-3169/a000367

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

DeBruine, L. M. (2005). Trustworthy but not lust-worthy: context-specific effects of facial resemblance. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1566), 919–922. https://doi.org/10.1098/RSPB.2004.3003

DeBruine, L. M., Jones, B. C., Little, A. C., & Perrett, D. I. (2008). Social perception of facial resemblance in humans. *Archives of Sexual Behavior*, *37*(1), 64–77. https://doi.org/10.1007/S10508-007-9266-0/TABLES/1 de Leersnyder, J., Mesquita, B., & Kim, H. S. (2011). Where do my emotions belong? a Study of immigrants' emotional acculturation. *Personality and Social Psychology Bulletin*, *37*(4), 451–463. https://doi.org/10.1177/0146167211399103

DeSteno, D., Breazeal, C., Frank, R. H., Pizarro, D., Baumann, J., Dickens, L., & Lee, J. J. (2012). Detecting the trustworthiness of novel partners in economic exchange. *Psychological Science*, *23*(12), 1549–1556. https://doi.org/10.1177/0956797612448793

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., Georgila, K., Gratch, J., Hartholt, A., Lhommet, M., Lucas, G., Marsella, S., Morbini, F., Nazarian, A., Scherer, S., Stratou, G., Suri, A., Traum, D., Wood, R., ... Morency, L. P. (2014). SimSensei kiosk: A virtual human interviewer for healthcare decision support. *13th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2014*, *2*, 1061–1068. http://simsensei.ict.usc.edu/

Dibeklioğlu, H., Salah, A. A., & Gevers, T. (2012). Are you really smiling at me? Spontaneous versus posed enjoyment smiles. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7574 LNCS*(PART 3), 525–538. https://doi.org/10.1007/978-3-642-33712-3_38

Diener, E., & Biswas-Diener, R. (2002). Will money increase subjective well-being? A literature review and guide to needed research. *Social Indicators Research*, *57*(2), 119–169. https://doi.org/10.1023/A:1014411319119

Diener, E., & Biswas-Diener, R. (2009). *Happiness: Unlocking the mysteries of psychological wealth*. Blackwell Publishing Ltd. https://doi.org/10.1002/9781444305159

Dotsch, R., Hassin, R. R., & Todorov, A. (2017). Statistical learning shapes face evaluation. *Nature Human Behaviour*, *1*(1), 0001. https://doi.org/10.1038/s41562-016-0001

Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science*, *3*(5), 562–571. https://doi.org/10.1177/1948550611430272

Drummond, P. D. (1994). The effect of anger and pleasure on facial blood flow. *Australian Journal of Psychology*, *46*(2), 95–99. https://doi.org/10.1080/00049539408259479

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, *42*(3-4), 177–190. https://doi.org/10.1016/S0921-8890(02)00374-3

Durante, F., Tablante, C. B., & Fiske, S. T. (2017). Poor but warm, rich but cold (and competent): Social classes in the Stereotype Content Model. *Journal of Social Issues*, *73*(1), 138–157. https://doi.org/10.1111/JOSI.12208

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences*, *98*(24), 13763–13768. https://doi.org/10.1073/PNAS.231499798

Eggleston, A., McCall, C., Cook, R., & Over, H. (2021). Parents reinforce the formation of first impressions in conversation with their children. *PLOS ONE*, *16*(8), e0256118. https://doi.org/10.1371/JOURNAL.PONE.0256118

Eisenegger, C., Haushofer, J., & Fehr, E. (2011). The role of testosterone in social interaction. *Trends in Cognitive Sciences*, *15*(6), 263–271. https://doi.org/10.1016/J.TICS.2011.04.008

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology II. *Journal of Personality and Social Psychology*, *58*(2), 342–353. https://doi.org/10.1037/0022-3514.58.2.342

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, *1*(1), 56–75. https://doi.org/10.1007/BF01115465

Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System: Investigator's guide*. Consulting Psychologists Press. https://books.google.co.uk/books/about/Facial_Action_Coding_System.html?id=7pqFtQAACAAJ%5C%5C%5C&redir_esc=y

Etcoff, N. L., Stock, S., Haley, L. E., Vickery, S. A., & House, D. M. (2011). Cosmetics as a feature of the extended human phenotype: Modulation of the perception of biologically important facial signals. *PLOS ONE*, *6*(10), e25656. https://doi.org/10.1371/JOURNAL.PONE.0025656

Farrier, C. E., Pearson, J. D., & Beran, T. N. (2020). Children's fear and pain during medical procedures: A quality improvement study with a humanoid robot. *The Canadian journal of nursing research = Revue canadienne de recherche en sciences infirmieres*, *52*(4), 328–334. https://doi.org/10.1177/0844562119862742

Fink, B., Matts, P. J., D'Emiliano, D., Bunse, L., Weege, B., & Röder, S. (2012). Colour homogeneity and visual perception of age, health and attractiveness of male facial skin. *Journal of the European Academy of Dermatology and Venereology*, *26*(12), 1486–1492. https://doi.org/10.1111/j.1468-3083.2011.04316.x

Fink, B., & Matts, P. (2008). The effects of skin colour distribution and topography cues on the perception of female facial age and health. *Journal of the European Academy of Dermatology and Venereology*, *22*(4), 493–498. https://doi.org/10.1111/j.1468-3083.2007.02512.x

Fink, B., Grammer, K., & Matts, P. J. (2006). Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evolution and Human Behavior*, *27*(6), 433–442. https://doi.org/10.1016/J.EVOLHUMBEHAV.2006.08.007

Fischer, A. H., Becker, D., & Veenstra, L. (2012). Emotional mimicry in social context: The case of disgust and pride. *Frontiers in Psychology*, *3*(NOV), 475. https://doi.org/10.3389/FPSYG.2012.00475/BIBTEX

Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, *27*(2), 67–73. https://doi.org/10.1177/0963721417738825

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

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

Fleischmann, A., Sieverding, M., Hespenheide, U., Weiß, M., & Koch, S. C. (2016). See feminine – Think incompetent? The effects of a feminine outfit on the evaluation of women's computer competence. *Computers & Education*, *95*, 63–74. https://doi.org/10.1016/J.COMPEDU.2015.12.007

Fletcher, G. J., Kerr, P. S., Li, N. P., & Valentine, K. A. (2014). Predicting romantic interest and decisions in the very early stages of mate selection: standards, accuracy, and sex differences. *Personality and Social Psychology Bulletin*, *40*(4), 540–550. https://doi.org/10.1177/0146167213519481

Foo, Y. Z., Sutherland, C. A., Burton, N. S., Nakagawa, S., & Rhodes, G. (2021). Accuracy in facial trustworthiness impressions: Kernel of truth or modern physiognomy. *Personality and Social Psychology Bulletin*. https://doi.org/10.1177/01461672211048110

Frank, M. G., Ekman, P., & Friesen, W. V. (1993). Behavioral markers and recognizability of the smile of enjoyment. *Journal of Personality and Social Psychology*, *64*(1), 83–93. https://doi.org/10.1037/0022-3514.64.1.83

Franklin, R. G., & Zebrowitz, L. A. (2016). The influence of political candidates' facial appearance on older and younger adults' voting choices and actual electoral success (T. Craig, Ed.). *Cogent Psychology*, *3*(1), 1151602. https://doi.org/10.1080/23311908.2016.1151602

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, *118*(2), 247–279. https://doi.org/10.1037/a0022327

Freeman, J. B., Stolier, R. M., & Brooks, J. A. (2020). Dynamic Interactive Theory as a domain-general account of social perception. *Advances in Experimental Social Psychology*, *61*, 237–287. https://nyuscholars.nyu.edu/en/publications/dynamic-interactive-theory-as-a-domain-general-account-of-social-%20https://doi.org/10.1016/bs.aesp.2019.09.005

Friedman, H., & Zebrowitz, L. A. (1992). The contribution of typical sex differences in facial maturity to sex role stereotypes. *Personality and Social Psychology Bulletin*, *18*(4), 430–438. https://doi.org/10.1177/0146167292184006

Geniole, S. N., Molnar, D. S., Carré, J. M., & McCormick, C. M. (2014). The facial width-to-height ratio shares stronger links with judgments of aggression than with judgments of trustworthiness. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1526–1541. https://doi.org/10.1037/a0036732

Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin. https://doi.org/10.4324/9781315740218

Gifford, R. (1994). A lens-mapping framework for understanding the encoding and decoding of interpersonal dispositions in nonverbal behavior. *Journal of Personality and Social Psychology*, *66*(2), 398–412. https://doi.org/10.1037/0022-3514.66.2.398

Gill, D., Garrod, O. G. B., Jack, R. E., & Schyns, P. G. (2014). Facial movements strategically camouflage involuntary social signals of face morphology. *Psychological Science*, *25*(5), 1079–1086. https://doi.org/10.1177/0956797614522274
Jack, Rachael/A-4692-2012 Jack, Rachael/0000-0003-3687-0799 14

Glikson, E., Cheshin, A., & van Kleef, G. A. (2018). The dark side of a smiley: Effects of smiling emoticons on virtual first impressions. *Social Psychological and Personality Science*, *9*(5), 614–625. https://doi.org/10.1177/1948550617720269

Gnisci, A., & Pace, A. (2014). The effects of hand gestures on psychosocial perception: A preliminary study. *Smart Innovation, Systems and Technologies*, *26*, 305–312. https://doi.org/10.1007/978-3-319-04129-2_30

Goldenthal, P., Johnston, R. E., & Kraut, R. E. (1981). Smiling, appeasement, and the silent bared-teeth display. *Ethology and Sociobiology*, *2*(3), 127–133. https://doi.org/10.1016/0162-3095(81)90025-X

Goodwin, G. P. (2015). Moral Character in Person Perception. *Current Directions in Psychological Science*, *24*(1), 38–44. https://doi.org/10.1177/0963721414550709

Graham, J. R., Harvey, C. R., & Puri, M. (2012). A corporate beauty contest. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1571469

Gray-Roncal, K., Fitzgerald, K. C., Ryerson, L. Z., Charvet, L., Cassard, S. D., Naismith, R., Ontaneda, D., Mahajan, K., Castro-Borrero, W., & Mowry, E. M. (2021). Association of disease severity and socioeconomic status in black and white americans with multiple sclerosis. *Neurology*, *97*(9), e881–e889. https://doi.org/10.1212/WNL.0000000000012362

Grigoryan, L., Bai, X., Durante, F., Fiske, S. T., Fabrykant, M., Hakobjanyan, A., Javakhishvili, N., Kadirov, K., Kotova, M., Makashvili, A., Maloku, E., Morozova-Larina, O., Mullabaeva, N., Samekin, A., Verbilovich, V., & Yahiiaiev, I. (2020). Stereotypes as historical accidents: Images of social class in postcommunist versus capitalist societies. *Personality and Social Psychology Bulletin*, *46*(6), 927–943. https://doi.org/10.1177/0146167219881434

Hall, E. T. . (1966). *The hidden dimension* ([1st ed.].). Doubleday.

Hareli, S., Shomrat, N., & Hess, U. (2009). Emotional versus neutral expressions and perceptions of social dominance and submissiveness. *Emotion*, *9*(3), 378–384. https://doi.org/10.1037/a0015958

Hareli, S., Smoly, M., & Hess, U. (2018). Help me Obi-Wan: the influence of facial dominance on perceptions of helpfulness. *Social Influence*, *13*(3), 163–176. https://doi.org/10.1080/15534510.2018.1500944

Hareli, S., Vider, E., & Hanoch, Y. (2021). The influence of facial dominance on perceptions of risk-taking preferences. *Basic and Applied Social Psychology*, *43*(5), 283–306. https://doi.org/10.1080/01973533.2021.1929988

Harris, C. R., & Alvarado, N. (2005). Facial expressions, smile types, and self-report during humour, tickle, and pain. *Cognition and Emotion*, *19*(5), 655–669. https://doi.org/10.1080/02699930441000472

Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: facial structure predicts unethical behaviour. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1728). https://doi.org/10.1098/RSPB.2011.1193

Haselhuhn, M. P., Wong, E. M., & Ormiston, M. E. (2013). Self-fulfilling prophecies as a link between men's facial Width-to-Height ratio and behavior. *PLoS ONE*, *8*(8), e72259. https://doi.org/10.1371/journal.pone.0072259

Hassin, R., & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology*, *78*(5), 837–852. https://doi.org/10.1037/0022-3514.78.5.837

Hayashi, N., Someya, N., Maruyama, T., Hirooka, Y., Endo, M. Y., & Fukuba, Y. (2009). Vascular responses to fear-induced stress in humans. *Physiology and Behavior*, *98*(4), 441–446. https://doi.org/10.1016/j.physbeh.2009.07.008

Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and dynamic facial cues differentially affect the consistency of social evaluations. *Personality and Social Psychology Bulletin*, *41*(8), 1123–1134. https://doi.org/10.1177/0146167215591495

Hehman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology*, *49*, 747–754. https://doi.org/10.1016/j.jesp.2013.02.015

Hehman, E., Sutherland, C. A., Flake, J. K., & Slepian, M. L. (2017). 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

Henderson, A. J., Holzleitner, I. J., Talamas, S. N., & Perrett, D. I. (2016). Perception of health from facial cues. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1693). https://doi.org/10.1098/rstb.2015.0380

Hess, U., Adams, R. B., Grammer, K., & Kleck, R. E. (2009). Face gender and emotion expression: Are angry women more like men? *Journal of Vision*, *9*(12), 19–19. https://doi.org/10.1167/9.12.19

Hess, U., Beaupré, M., & Cheung, N. (2002). Who to whom and why – cultural differences and similarities in the function of smiles. *An empirical reflection on the smile* (pp. 187–216). Edwin Mellen Press. https://www.researchgate.net/publication/222712479_Who_to_whom_and_why_-_cultural_differences_and_similarities_in_the_function_of_smiles

Hess, U., Blairy, S., & Kleck, R. E. (2000). The influence of facial emotion displays, gender, and ethnicity on judgments of dominance and affiliation. *Journal of Nonverbal Behavior*, *24*(4), 265–283. https://doi.org/10.1023/A:1006623213355

Hester, N., Jones, B. C., & Hehman, E. (2021). Perceived femininity and masculinity contribute independently to facial impressions. *Journal of Experimental Psychology: General*, *150*(6), 1147–1164. https://doi.org/10.1037/xge0000989

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*(2), 297–315. https://doi.org/10.1037/a0022128

Ibrahimagić-Šeper, L., Čelebić, A., Petričević, N., & Selimović, E. (2006). Anthropometric differences between males and females in face dimensions and dimensions of central maxillary incisors. *Medicinski Glasnik*, *3*(2), 58–62.

Imhoff, R., & Koch, A. (2017). How orthogonal are the big two of social perception? On the curvilinear relation between agency and communion. *Perspectives on Psychological Science*, *12*(1), 122–137. https://doi.org/10.1177/1745691616657334

Imhoff, R., Woelki, J., Hanke, S., & Dotsch, R. (2013). Warmth and competence in your face! Visual encoding of stereotype content. *Frontiers in Psychology*, *4*, 386. https://doi.org/10.3389/fpsyg.2013.00386

Ince, R. A. A., Paton, A. T., Kay, J. W., & Schyns, P. G. (2021). Bayesian inference of population prevalence. *eLife*, *10*. https://doi.org/10.1101/2020.07.08.191106

Jack, R. E., & Schyns, P. G. (2017). Toward a social psychophysics of face communication (S. T. Fiske, Ed.). *Annual Review of Psychology*, *68*, 269–297. https://doi.org/10.1146/annurev-psych-010416-044242

Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General*, *141*(1), 19–25. https://doi.org/10.1037/a0023463

Jack, R. E., Garrod, O. G., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology*, *24*(2), 187–192. https://doi.org/10.1016/j.cub.2013.11.064

Jack, R. E., Sun, W., Delis, I., Garrod, O. G. B., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion. *Journal of Experimental Psychology-General*, *145*(6), 708–730. https://doi.org/10.1037/xge0000162

Jacobus, S., & Loy, J. (1981). The grimace and gecker - A submissive display among Patas monkeys. *PRIMATES*, *22*(3), 393.

Jaeger, B., Evans, A. M., Stel, M., & van Beest, I. (2019). Explaining the persistent influence of facial cues in social decision-making. *Journal of Experimental Psychology: General*, *148*(6), 1008–1021. https://doi.org/10.1037/xge0000591

Jaeger, B., & Jones, A. L. (2022). Which facial features are central in impression formation? *Social Psychological and Personality Science*, *13*(2), 553–561. https://doi.org/10.1177/19485506211034979

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*. https://doi.org/10.1016/j.jesp.2020.104004

Jaeger, B., Wagemans, F. M. A., Evans, A. M., & van Beest, I. (2018). Effects of facial skin smoothness and blemishes on trait impressions. *Perception*, *47*(6), 608–625. https://doi.org/10.1177/0301006618767258

Janssens, S. E., Sack, A. T., Jessen, S., & de Graaf, T. A. (2020). Can processing of face trustworthiness bypass early visual cortex? A transcranial magnetic stimulation masking study. *Neuropsychologia*, *137*. https://doi.org/10.1016/j.neuropsychologia.2019.107304

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

Jones, A. L. (2018). The influence of shape and colour cue classes on facial health perception. *Evolution and Human Behavior*, *39*(1), 19–29. https://doi.org/10.1016/J.EVOLHUMBEHAV.2017.09.005

Jones, A. L., & Kramer, R. S. (2021). Facial first impressions form two clusters representing approach-avoidance. *Cognitive Psychology*, *126*, 101387. https://doi.org/10.1016/j.cogpsych.2021.101387

Jones, A. L., Kramer, R. S., & Ward, R. (2012). Signals of personality and health: The contributions of facial shape, skin texture, and viewing angle. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1353–1361. https://doi.org/10.1037/a0027078

Jones, B. C., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., Arinze, N. C., Ndukaihe, I. L. G., Bloxsom, N. G., Lewis, S. C., Foroni, F., Willis, M. L., Cubillas, C. P., Vadillo, M. A., Turiegano, E., Gilead, M., Simchon, A., Saribay, S. A., Owsley, N. C., Jang, C., . . . Coles, N. A. (2021). To which world regions does the valence–dominance model of social perception apply? *Nature Human Behaviour*, *5*(1), 159–169. https://doi.org/10.1038/s41562-020-01007-2

Joo, J., Steen, F. F., & Zhu, S. C. (2015). Automated facial trait judgment and election outcome prediction: Social dimensions of face. *Proceedings of the IEEE International*

Conference on Computer Vision, *2015 Inter*, 3712–3720. https://doi.org/10.1109/ICCV.2015.423

Judd, C. M., Garcia-Marques, T., & Yzerbyt, V. Y. (2019). The complexity of relations between dimensions of social perception: Decomposing bivariate associations with crossed random factors. *Journal of Experimental Social Psychology*, *82*, 200–207. https://doi.org/10.1016/J.JESP.2019.01.008

Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, *89*(6), 899–913. https://doi.org/10.1037/0022-3514.89.6.899

Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization. *Advances in experimental social psychology* (pp. 1–80). Academic Press. https://doi.org/10.1016/bs.aesp.2016.10.001

Keles, U., Lin, C., & Adolphs, R. (2021). A cautionary note on predicting social judgments from faces with deep neural networks. *Affective Science*, *2*(4), 438–454. https://doi.org/10.1007/S42761-021-00075-5

Keltner, D. (1995). Signs of appeasement - Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology*, *68*(3), 441–454. https://doi.org/10.1037/0022-3514.68.3.441

Keltner, D., Young, R. C., & Buswell, B. N. (1997). Appeasement in human emotion, social practice, and personality. *Aggressive Behavior*, *23*(5), 359–374. https://doi.org/10.1002/(SICI)1098-2337(1997)23:5<359::AID-AB5>3.0.CO;2-D

Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review*, *98*(2), 155–163. https://doi.org/10.1037/0033-295X.98.2.155

Kenny, P., Hartholt, A., Gratch, J., Swartout, W., Traum, D., Marsella, S., Piepol, D., & Rey, M. D. (2007). Building interactive virtual humans for training environments. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, (7105), 1–16.

Keres, A., & Chartier, C. R. (2016). The biasing effects of visual background on perceived facial trustworthiness. *Psi Chi Journal of Psychological Research*, *21*(3), 170–175. https://doi.org/10.24839/b21.3.170

Kervyn, N., Fiske, S., & Yzerbyt, V. (2015). Forecasting the primary dimension of social perception: Symbolic and realistic threats together predict warmth in the Stereotype Content Model. *Social psychology*, *46*(1), 36. https://doi.org/10.1027/1864-9335/A000219

Ketelaar, T., Koenig, B. L., Gambacorta, D., Dolgov, I., Hor, D., Zarzosa, J., Luna-Nevarez, C., Klungle, M., & Wells, L. (2012). Smiles as signals of lower status in football players and fashion models: Evidence that smiles are associated with lower dominance and lower prestige. *Evolutionary Psychology*, *10*(3), 371–397. https://doi.org/10.1177/147470491201000301

Klapper, A., Dotsch, R., van Rooij, I., & Wigboldus, D. H. (2016). Do we spontaneously form stable trustworthiness impressions from facial appearance? *Journal of Personality and Social Psychology*, *111*(5), 655–664. https://doi.org/10.1037/PSPA0000062

Klatt, J., Eimler, S. C., & Krämer, N. C. (2016). Makeup your mind: The impact of styling on perceived competence and warmth of female leaders. *The Journal of Social Psychology*, *156*(5), 483–497. https://doi.org/10.1080/00224545.2015.1129303

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16. https://nyuscholars.nyu.edu/en/publications/whats-new-in-psychtoolbox-3

Kleisner, K., Kočnar, T., Rubešová, A., & Flegr, J. (2010). Eye color predicts but does not directly influence perceived dominance in men. *Personality and Individual Differences*, *49*(1), 59–64. https://doi.org/10.1016/J.PAID.2010.03.011

Kleisner, K., Priplatova, L., Frost, P., & Flegr, J. (2013). Trustworthy-looking face meets brown eyes (D. G. Pelli, Ed.). *PLoS ONE*, *8*(1), e53285. https://doi.org/10.1371/journal.pone.0053285

Klofstad, C. A. (2017). Looks and sounds like a winner: Perceptions of competence in candidates' faces and voices influences vote choice. *Journal of Experimental Political Science*, *4*(3), 229–240. https://doi.org/10.1017/XPS.2017.19

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior*, *20*(3), 165–182. https://doi.org/10.1007/BF02281954

Koppensteiner, M., Stephan, P., & Jäschke, J. P. M. (2016). Moving speeches: Dominance, trustworthiness and competence in body motion. *Personality and Individual Differences*, *94*, 101–106. https://doi.org/10.1016/j.paid.2016.01.013

Kraus, M. W., Tan, J. J., & Tannenbaum, M. B. (2013). The social ladder: A rank-based perspective on social class. *Psychological Inquiry*, *24*(2), 81–96. https://doi.org/10.1080/1047840X.2013.778803

Kraus, M. W., Torrez, B., Park, J. W., & Ghayebi, F. (2019). Evidence for the reproduction of social class in brief speech. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(46), 22998–23003. https://doi.org/10.1073/PNAS.1900500116

Kreibig, S. D., Wilhelm, F. H., Roth, W. T., & Gross, J. J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films. *Psychophysiology*, *44*(5), 787–806. https://doi.org/10.1111/j.1469-8986.2007.00550.x

Krumhuber, E. G., & Manstead, A. S. (2009). Can Duchenne smiles be feigned? New evidence on felt and false smiles. *Emotion*, *9*(6), 807–820. https://doi.org/10.1037/A0017844

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

Kunz, M., Prkachin, K., & Lautenbacher, S. (2009). The smile of pain. *Pain*, *145*(3), 273–275. https://doi.org/10.1016/J.PAIN.2009.04.009

Kyle, D. J., & Mahler, H. I. M. (1996). The effects of hair color and cosmetic use on perceptions of a female's ability. *Psychology of Women Quarterly*, *20*, 447–455.

Lavan, N., Mileva, M., Burton, A. M., Young, A. W., & McGettigan, C. (2021). Trait evaluations of faces and voices: Comparing within- and between-person variability. *Journal of Experimental Psychology: General*, *150*(9), 1854–1869. https://doi.org/10.1037/xge0001019

Lavan, N., Mileva, M., & McGettigan, C. (2021). How does familiarity with a voice affect trait judgements? *British Journal of Psychology*, *112*(1), 282–300. https://doi.org/10.1111/bjop.12454

Lee, K. M., Peng, W., Jin, S.-A., & Yan, C. (2006). Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of Communication*, *56*(4), 754–772. https://doi.org/10.1111/j.1460-2466.2006.00318.x

Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. (2013). Telling facial metrics: facial width is associated with testosterone levels in men. *Evolution and Human Behavior*, *34*(4), 273–279. https://doi.org/10.1016/J.EVOLHUMBEHAV.2013.03.005

Lei, R. F., & Bodenhausen, G. V. (2017). Racial assumptions color the mental representation of social class. *Frontiers in Psychology*, *8*(MAR), 519. https://doi.org/10.3389/FPSYG.2017.00519/BIBTEX

Levenson, R. W. (2003). Blood, sweat, and fears: The autonomic architecture of emotion. *Annals of the New York Academy of Sciences*, *1000*, 348–366. https://doi.org/10.1196/annals.1280.016

Lin, C., Keles, U., & Adolphs, R. (2021). Four dimensions characterize attributions from faces using a representative set of English trait words. *Nature Communications*, *12*(1), 1–15. https://doi.org/10.1038/s41467-021-25500-y

Lin, M. H., Kwan, V. S., Cheung, A., & Fiske, S. T. (2005). Stereotype content model explains prejudice for an envied outgroup: Scale of anti-Asian American stereotypes. *Personality and Social Psychology Bulletin*, *31*(1), 34–47. https://doi.org/10.1177/0146167204271320

Linke, L., Saribay, S. A., & Kleisner, K. (2016). Perceived trustworthiness is associated with position in a corporate hierarchy. *Personality and Individual Differences*, *99*, 22–27. https://doi.org/10.1016/J.PAID.2016.04.076

Little, A. C. (2014). Facial attractiveness. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(6), 621–634. https://doi.org/10.1002/WCS.1316

Little, A. C., Roberts, S. C., Jones, B. C., & DeBruine, L. M. (2012). The perception of attractiveness and trustworthiness in male faces affects hypothetical voting decisions differently in wartime and peacetime scenarios. *Quarterly Journal of Experimental Psychology*, *65*(10), 2018–2032. https://doi.org/10.1080/17470218.2012.677048

Luttrell, A., Sacchi, S., & Brambilla, M. (2022). Changing impressions in competence-oriented domains: The primacy of morality endures. *Journal of Experimental Social Psychology*, *98*, 104246. https://doi.org/10.1016/j.jesp.2021.104246

Lyons, M., & Simeonov, L. (2016). The undesirable Dark Triad? Women dislike Dark Triad male faces across different mating context and socio-ecological conditions. *Personality and Individual Differences*, *90*, 338–341. https://doi.org/10.1016/j.paid.2015.11.047

Macrae, C. N., & Martin, D. (2007). A boy primed Sue: Feature-based processing and person construal. *European Journal of Social Psychology*, *37*(5), 793–805. https://doi.org/10.1002/ejsp.406

Main, J. C., Jones, B. C., DeBruine, L. M., & Little, A. C. (2009). Integrating gaze direction and sexual dimorphism of face shape when perceiving the dominance of others. *Perception*, *38*(9), 1275–1283. https://doi.org/10.1068/P6347

Malatesta, C. Z., Fiore, M. J., & Messina, J. J. (1987). Affect, personality, and facial expressive characteristics of older people. *Psychology and aging*, *2*(1), 64–69. https://doi.org/10.1037/0882-7974.2.1.64

Maricchiolo, F., Gnisci, A., Bonaiuto, M., & Ficca, G. (2009). Effects of different types of hand gestures in persuasive speech on receivers' evaluations. *Language and Cognitive Processes*, *24*(2), 239–266. https://doi.org/10.1080/01690960802159929

Marmot, M. G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., Feeney, A., Marmot, M. G., & Smith, G. D. (1991). Health inequalities among British civil servants: the Whitehall II study. *Lancet*, *337*(8754), 1387–1393. https://doi.org/10.1016/0140-6736(91)93068-K

Marsella, S., & Gratch, J. (2003). Modeling coping behavior in virtual humans: Don't worry, be happy. *Proceedings of the International Conference on Autonomous Agents*, *2*, 313–320. https://doi.org/10.1145/860575.860626

Martin, A. E., & Slepian, M. L. (2021). The primacy of gender: Gendered cognition underlies the Big Two dimensions of social cognition. *Perspectives on Psychological Science*, *16*(6), 1143–1158. https://doi.org/10.1177/1745691620904961

Martin, J., Rychlowska, M., Wood, A., & Niedenthal, P. (2017). Smiles as multipurpose social signals. *Trends in Cognitive Sciences*, *21*(11), 864–877. https://doi.org/10.1016/j.tics.2017.08.007

Masi, M., Mattavelli, S., Fasoli, F., & Brambilla, M. (2021). Cross-modal impression updating: Dynamic impression updating from face to voice and the other way around. *British Journal of Social Psychology*. https://doi.org/10.1111/bjso.12511

Mattarozzi, K., Todorov, A., Marzocchi, M., Vicari, A., & Russo, P. M. (2015). Effects of gender and personality on first impression (T. Zalla, Ed.). *PLoS ONE*, *10*(9), 1–13. https://doi.org/10.1371/journal.pone.0135529

McAleer, P., Todorov, A., & Belin, P. (2014). How do you say 'hello'? Personality impressions from brief novel voices. *PLOS ONE*, *9*(3), e90779. https://doi.org/10.1371/journal.pone.0090779

McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, *90*(3), 215–238. https://doi.org/10.1037/0033-295X.90.3.215

McArthur, L. Z., & Apatow, K. (1984). Impressions of baby-faced adults. *Social Cognition*, *2*(4), 315–342. https://doi.org/10.1521/soco.1984.2.4.315

McDuff, D. (2016). Discovering facial expressions for states of amused, persuaded, informed, sentimental and inspired. *ICMI 2016 - Proceedings of the 18th ACM International Conference on Multimodal Interaction*, 71–75. https://doi.org/10.1145/2993148.2993192

Mignault, A., & Chaudhuri, A. (2003). The many faces of a neutral face: Head tilt and perception of dominance and emotion. *Journal of Nonverbal Behavior 2003 27:2*, *27*(2), 111–132. https://doi.org/10.1023/A:1023914509763

Mileva, M., & Lavan, N. (2022). How quickly can we form a trait impression from voices? *PsyArXiv*. https://doi.org/10.31234/OSF.IO/ZD4UN

Miller, J. K., Westerman, D. L., & Lloyd, M. E. (2004). Are first impressions lasting impressions? An exploration of the generality of the primacy effect in memory for repetitions. *Memory and Cognition*, *32*(8), 1305–1315. https://doi.org/10.3758/BF03206321

Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior*, *27*(4), 237–254. https://doi.org/10.1023/A:1027332800296

Morrison, A. (2019). Contributive justice: social class and graduate employment in the UK. *https://doi.org/10.1080/13639080.2019.1646414*, *32*(4), 335–346. https://doi.org/10.1080/13639080.2019.1646414

Mueller, U., & Mazur, A. (1996). Facial dominance of West Point cadets as a predictor of later military rank. *Social Forces*, *74*(3), 823–850. https://doi.org/10.1093/sf/74.3.823

Nauts, S., Langner, O., Huijsmans, I., Vonk, R., & Wigboldus, D. H. (2014). Forming impressions of personality: A replication and review of Asch's (1946) evidence for a primacy-of-warmth effect in impression formation. *Social Psychology*, *45*(3), 153–163. https://doi.org/10.1027/1864-9335/a000179

Neave, N., & Shields, K. (2008). The effects of facial hair manipulation on female perceptions of attractiveness, masculinity, and dominance in male faces. *Personality and Individual Differences*, *45*(5), 373–377. https://doi.org/10.1016/J.PAID.2008.05.007

Nelissen, R. M., & Meijers, M. H. (2011). Social benefits of luxury brands as costly signals of wealth and status. *Evolution and Human Behavior*, *32*(5), 343–355. https://doi.org/10.1016/J.EVOLHUMBEHAV.2010.12.002

Nguyen, L. S., Marcos-Ramiro, A., Marrón Romera, M., & Gatica-Perez, D. (2013). Multimodal analysis of body communication cues in employment interviews. *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*, 437–444. https://doi.org/10.1145/2522848.2522860

Niedenthal, P. M., Mermillod, M., Maringer, M., & Hess, U. (2010). The Simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behavioral and Brain Sciences*, *33*(6), 417–433. https://doi.org/10.1017/S0140525X10000865

Nishi, A., Shirado, H., Rand, D. G., & Christakis, N. A. (2015). Inequality and visibility of wealth in experimental social networks. *Nature*, *526*(7573), 426–429. https://doi.org/10.1038/nature15392

North, M. S., & Fiske, S. T. (2015). Modern attitudes toward older adults in the aging world: A cross-cultural meta-analysis. *Psychological Bulletin*, *141*(5), 993–1021. https://doi.org/10.1037/a0039469

Oh, D. W., Buck, E. A., & Todorov, A. (2019). Revealing hidden gender biases in competence impressions of faces. *Psychological Science*, *30*(1), 65–79. https://doi.org/10.1177/0956797618813092

Ohala, J. J. (1982). The voice of dominance. *The Journal of the Acoustical Society of America*, *72*(S1), S66–S66. https://doi.org/10.1121/1.2020007

Oliveira, M., Garcia-Marques, T., Dotsch, R., & Garcia-Marques, L. (2019). Dominance and competence face to face: Dissociations obtained with a reverse correlation approach. *European Journal of Social Psychology*, *49*(5), 888–902. https://doi.org/10.1002/ejsp.2569

Olivola, C. Y., Eubanks, D. L., & Lovelace, J. B. (2014). The many (distinctive) faces of leadership: Inferring leadership domain from facial appearance. *The Leadership Quarterly*, *25*(5), 817–834. https://doi.org/10.1016/J.LEAQUA.2014.06.002

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

O'Neil, S. F., & Webster, M. A. (2011). Adaptation and the perception of facial age. *Visual Cognition*, *19*(4), 534–550. https://doi.org/10.1080/13506285.2011.561262

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(32), 11087–92. 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

Otta, E., Lira, B. B. P., Delevati, N. M., Cesar, O. P., & Pires, C. S. G. (1994). The effect of smiling and of head tilting on person perception. *The Journal of psychology*, *128*(3), 323–331. https://doi.org/10.1080/00223980.1994.9712736

Pace, A., & Gnisci, A. (2019). Some gestures are better than others: The warmth and competence effect of hand gestures in interaction. *Psicologia Sociale*, *14*(2), 281–302. https://doi.org/10.1482/94269

Penton-Voak, I. S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior*, *25*(4), 229–241. https://doi.org/10.1016/J.EVOLHUMBEHAV.2004.04.003

Philip, P., Micoulaud-Franchi, J. A., Sagaspe, P., Sevin, E. D., Olive, J., Bioulac, S., & Sauteraud, A. (2017). Virtual human as a new diagnostic tool, a proof of concept study in the field of major depressive disorders. *Scientific Reports*, *7*(1), 1–7. https://doi.org/10.1038/srep42656

Pivonkova, V., Rubesova, A., Lindova, J., & Havlicek, J. (2011). Sexual dimorphism and personality attributions of male faces. *Archives of Sexual Behavior*, *40*(6), 1137–1143. https://doi.org/10.1007/S10508-011-9821-6/TABLES/2

Ponsot, E., Arias, P., & Aucouturier, J.-J. (2018). Uncovering mental representations of smiled speech using reverse correlation. *The Journal of the Acoustical Society of America*, *143*(1), EL19–EL24. https://doi.org/10.1121/1.5020989

Porter, S., England, L., Juodis, M., Ten Brinke, L., & Wilson, K. (2008). Is the face a window to the soul? Investigation of the accuracy of intuitive judgments of the trustworthiness of human faces. *Canadian Journal of Behavioural Science*, *40*(3), 171–177. https://doi.org/10.1037/0008-400X.40.3.171

Poutvaara, P., Jordahl, H., & Berggren, N. (2009). Faces of politicians: Babyfacedness predicts inferred competence but not electoral success. *Journal of Experimental Social Psychology*, *45*(5), 1132–1135. https://doi.org/10.1016/J.JESP.2009.06.007

Prakash, A., & Rogers, W. A. (2016). Others: Effects of human-likeness and task. *International Journal of Social Robotics*, *7*(2), 309–331. https://doi.org/10.1007/s12369-014-0269-4.Why

Prkachin, K. M. (2009). Assessing pain by facial expression: Facial expression as nexus. *Pain Research and Management*, *14*(1), 53–58. https://doi.org/10.1155/2009/542964

Prkachin, K. M., & Solomon, P. E. (2008). The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain. *Pain*, *139*(2), 267–274. https://doi.org/10.1016/J.PAIN.2008.04.010

Puts, D. A., Gaulin, S. J., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, *27*(4), 283–296. https://doi.org/10.1016/j.evolhumbehav.2005.11.003

Qi, Y., Luo, Y., Feng, Y., Chen, Z., Li, Q., Du, F., & Liu, X. (2021). Trustworthy faces make people more risk-tolerant: The effect of facial trustworthiness on risk decision-making under gain and loss conditions. *PsyCh Journal*, *11*(1), 43–50. https://doi.org/10.1002/pchj.498

Quadflieg, S., & Westmoreland, K. (2019). Making sense of other people's encounters: Towards an integrative model of relational impression formation. *Journal of Nonverbal Behavior*, *43*(2), 233–256. https://doi.org/10.1007/s10919-019-00295-1

Re, D. E., & Rule, N. O. (2016). Making a (false) impression: The role of business experience in first impressions of CEO leadership ability. *Journal of Nonverbal Behavior*, *40*(3), 235–245. https://doi.org/10.1007/s10919-016-0231-2

Reiss, F., Meyrose, A.-K., Otto, C., Lampert, T., Klasen, F., & Ravens-Sieberer, U. (2019). Socioeconomic status, stressful life situations and mental health problems in children and adolescents: Results of the German BELLA cohort-study. *PLOS ONE*, *14*(3), e0213700. https://doi.org/10.1371/JOURNAL.PONE.0213700

Rennung, M., Blum, J., & Göritz, A. S. (2016). To strike a pose: No stereotype backlash for power posing women. *Frontiers in Psychology*, *7*(SEP), 1463. https://doi.org/10.3389/fpsyg.2016.01463

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

Rezlescu, C., Penton, T., Walsh, V., Tsujimura, H., Scott, S. K., & Banissy, M. J. (2015). Dominant Voices and Attractive Faces: The Contribution of Visual and Auditory Information to Integrated Person Impressions. *Journal of Nonverbal Behavior*, *39*(4), 355–370. https://doi.org/10.1007/s10919-015-0214-8

Richardson, J. T. E., Mittelmeier, J., & Rienties, B. (2020). The role of gender, social class and ethnicity in participation and academic attainment in UK higher education: an update. *https://doi.org/10.1080/03054985.2019.1702012*, *46*(3), 346–362. https://doi.org/10.1080/03054985.2019.1702012

Richey, M. H., Mcclelland, L., & Shimkunas, A. M. (1967). Relative influence of positive and negative information in impression formation and persistence. *Journal of Personality and Social Psychology*, *6*(3), 322–327. https://doi.org/10.1037/h0024734

Ritchie, K. L., Palermo, R., & Rhodes, G. (2017). Forming impressions of facial attractiveness is mandatory. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-00526-9

Rivera, L. A. (2012). Hiring as cultural matching: The case of elite professional service firms. *https://doi.org/10.1177/0003122412463213*, *77*(6), 999–1022. https://doi.org/10.1177/0003122412463213

Rivera, L. A., & Tilcsik, A. (2016). Class Advantage, Commitment Penalty: The Gendered Effect of Social Class Signals in an Elite Labor Market. *American Sociological Review*, *81*(6), 1097–1131. https://doi.org/10.1177/0003122416668154

Rizzo, A., Lange, B., Buckwalter, J. G., Forbell, E., Kim, J., Sagae, K., Williams, J., Difede, J. A., Rothbaum, B. O., Reger, G., Parsons, T., & Kenny, P. (2011). SimCoach: An intelligent virtual human system for providing healthcare information and support. *International Journal on Disability and Human Development*, *10*(4), 277–281. https://doi.org/10.1515/IJDHD.2011.046/MACHINEREADABLECITATION/RIS

Rohrmann, S., & Hopp, H. (2008). Cardiovascular indicators of disgust. *International Journal of Psychophysiology*, *68*(3), 201–208. https://doi.org/10.1016/j.ijpsycho.2008.01.011

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, *9*(4), 283–294. https://doi.org/10.1037/h0026086

Rosenthal, A. (2004). Raising hair. *Eighteenth-Century Studies*, *38*(1), 1–16. https://doi.org/10.1353/ecs.2004.0064

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*(100), 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Rudert, S. C., Reutner, L., Greifeneder, R., & Walker, M. (2017). Faced with exclusion: Perceived facial warmth and competence influence moral judgments of social exclusion. *Journal of Experimental Social Psychology*, *68*, 101–112. https://doi.org/10.1016/J.JESP.2016.06.005

Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward Agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology*, *77*(5), 1004–1010. https://doi.org/10.1037/0022-3514.77.5.1004

Rule, N. O., Adams, R. B., Ambady, N., & Freeman, J. B. (2012). Perceptions of dominance following glimpses of faces and bodies. *Perception*, *41*(6), 687–706. https://doi.org/10.1068/p7023

Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal* *of Personality and Social Psychology*, *104*(3), 409–426. https://doi.org/10.1037/a0031050

Russell, R., Porcheron, A., Sweda, J. R., Jones, A. L., Mauger, E., & Morizot, F. (2016). Facial contrast is a cue for perceiving health from the face. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(9), 1354–1362. https://doi.org/10.1037/xhp0000219

Rychlowska, M., Jack, R. E., Garrod, O. G., Schyns, P. G., Martin, J. D., & Niedenthal, P. M. (2017). Functional smiles: Tools for love, sympathy, and war. *Psychological Science*, *28*(9), 1259–1270. https://doi.org/10.1177/0956797617706082

Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*, *9*(2), 260–264. https://doi.org/10.1037/a0014681

Said, C. P., & Todorov, A. (2011). A statistical model of facial attractiveness. *Psychological science*, *22*(9), 1183–1190. https://doi.org/10.1177/0956797611419169

Samson, N., Fink, B., Matts, P. J., Dawes, N. C., & Weitz, S. (2010). Visible changes of female facial skin surface topography in relation to age and attractiveness perception. *Journal of Cosmetic Dermatology*, *9*(2), 79–88. https://doi.org/10.1111/j.1473-2165.2010.00489.x

Sapolsky, R. M. (2004). Social status and health in humans and other animals. *Annual Review of Anthropology*, *33*, 393–418. https://doi.org/10.1146/annurev.anthro.33.070203.144000

Satchell, L. P., Mayes, H. S., Lee, A. J., O'Reilly, L. C., Akehurst, L., & Morris, P. (2021). Is threat in the way they move? Influences of static and gait information on threat judgments of unknown people. *Evolutionary Psychological Science*, *7*(1), 61–68. https://doi.org/10.1007/S40806-020-00249-7/TABLES/1

Scherer, K. R. (1978). Personality inference from voice quality: The loud voice of extroversion. *European Journal of Social Psychology*, *8*(4), 467–487. https://doi.org/10.1002/ejsp.2420080405

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, *40*(1-2), 227–256. https://doi.org/10.1016/S0167-6393(02)00084-5

Schofield, T. P., Suomi, A., & Butterworth, P. (2021). Is the stereotype of welfare recipients associated with type of welfare state regime? A cross-national meta-regression of the stereotype content model. *Journal of Applied Social Psychology*. https://doi.org/10.1111/jasp.12843

Schuring, M., Robroek, S. J., Otten, F. W., Arts, C. H., & Burdorf, A. (2013). The effect of ill health and socioeconomic status on labor force exit and re-employment: A prospective study with ten years follow-up in the Netherlands. *Scandinavian Journal of Work, Environment and Health*, *39*(2), 134–143. https://doi.org/10.5271/sjweh.3321

Schyns, P. G., Goldstone, R. L., & Thibaut, J. P. (1998). The development of features in object concepts. *The Behavioral and brain sciences*, *21*(1), 1–54. https://doi.org/10.1017/S0140525X98000107

Shahraki, M., Agheli, L., Assari Arani, A., Sadeqi, H., & Ghaderi, S. (2018). The relationship between mother's socioeconomic status and child health. *Journal of Research and Health*, *8*(2), 143–151. https://doi.org/10.29252/JRH.8.2.143

Shang, J., & Li, Y. (2020). The effects of participants' sex and the facial trustworthiness of proposers on third-party decision-making in a dictator game. *PsyCh Journal*, *9*(6), 877–884. https://doi.org/10.1002/pchj.388

Shannon, C. E. (1948). *A Mathematical theory of communication* (tech. rep. No. 3). https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shen, X., & Ferguson, M. J. (2021). How resistant are implicit impressions of facial trustworthiness? When new evidence leads to durable updating. *Journal of Experimental Social Psychology*, *97*, 104219. https://doi.org/10.1016/j.jesp.2021.104219

Shields, S. A., MacArthur, H. J., & McCormick, K. T. (2017). The gendering of emotion and the psychology of women. *APA handbook of the psychology of women: History, theory, and battlegrounds*, *1*, 189–206. https://doi.org/10.1037/0000059-010

Shimbo, D., Rosenberg, L. B., Chaplin, W., Zhao, S., Goldensohn, E. R., Cholankeril, M., Fu, J., Hong, S. B., Jelic, S., & Burg, M. M. (2013). Endothelial cell activation, reduced endothelial cell reparative capacity, and impaired endothelial-dependent vasodilation after anger provocation. *International Journal of Cardiology*, *167*(3), 1064–1065. https://doi.org/10.1016/j.ijcard.2012.10.069

Singh, A., & Manjaly, J. A. (2021). The distress smile and its cognitive antecedents. *Journal of Nonverbal Behavior*, *45*(1), 11–30. https://doi.org/10.1007/S10919-020-00345-Z/FIGURES/1

Slepian, M. L., & Ames, D. R. (2016). Internalized impressions: The link between apparent facial trustworthiness and deceptive behavior is mediated by targets' expectations of how they will be judged. *Psychological Science*, *27*(2), 282–288. https://doi.org/10.1177/0956797615594897

Slepian, M. L., Young, S. G., & Harmon-Jones, E. (2017). An approach-avoidance motivational model of trustworthiness judgments. *Motivation Science*, *3*(1), 91–97. https://doi.org/10.1037/mot0000046

Smith, H. M., Dunn, A. K., Baguley, T., & Stacey, P. C. (2016). Concordant cues in faces and voices: Testing the Backup Signal Hypothesis. *Evolutionary Psychology*, *14*(1), 147470491663031. https://doi.org/10.1177/1474704916630317

Song, H., Vonasch, A. J., Meier, B. P., & Bargh, J. A. (2012). Brighten up: Smiles facilitate perceptual judgment of facial lightness. *Journal of Experimental Social Psychology*, *48*(1), 450–452. https://doi.org/10.1016/J.JESP.2011.10.003

Song, S., Jaiswal, S., Sanchez, E., Tzimiropoulos, G., Shen, L., & Valstar, M. (2021). Self-supervised learning of person-specific facial dynamics for automatic personality recognition. *IEEE Transactions on Affective Computing*. https://doi.org/10.1109/TAFFC.2021.3064601

South Palomares, J. K., & Young, A. W. (2018). Facial first impressions of partner preference traits: Trustworthiness, status, and attractiveness. *Social Psychological and Personality Science*, *9*(8), 990–1000. https://doi.org/10.1177/1948550617732388

Spencer, B., & Castano, E. (2007). Social class is dead. Long live social class! Stereotype threat among low socioeconomic status individuals. *Social Justice Research*, *20*(4), 418–432. https://doi.org/10.1007/S11211-007-0047-7

Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(19), 7710–7715. https://doi.org/10.1073/PNAS.1014345108/-/DCSUPPLEMENTAL

Stephen, I. D., Oldham, F. H., Perrett, D. I., & Barton, R. A. (2012). Redness enhances perceived aggression, dominance and attractiveness in men's faces. *Evolutionary Psychology*, *10*(3), 562–572. https://doi.org/10.1177/147470491201000312

Stins, J. F., Roelofs, K., Villan, J., Kooijman, K., Hagenaars, M. A., & Beek, P. J. (2011). Walk to me when I smile, step back when I'm angry: emotional faces modulate whole-body approach–avoidance behaviors. *Experimental Brain Research*, *212*(4), 603–611. https://doi.org/10.1007/S00221-011-2767-Z

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science*, *21*(3), 349–354. https://doi.org/10.1177/0956797610362647

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. (2018). A dynamic structure of social trait space. *Trends in cognitive sciences*, *22*(3), 197–200. https://doi.org/10.1016/j.tics.2017.12.003

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

Sung, K. T. (2001). Elder respect: exploration of ideals and forms in East Asia. *Journal of Aging Studies*, *15*(1), 13–26. https://doi.org/10.1016/S0890-4065(00)00014-1

Sung, K. T. (2004). Elder respect among young adults: A cross-cultural study of Americans and Koreans. *Journal of Aging Studies*, *18*(2), 215–230. https://doi.org/10.1016/J.JAGING.2004.01.002

Sussman, A. B., Petkova, K., & Todorov, A. (2013). Competence ratings in US predict presidential election outcomes in Bulgaria. *Journal of Experimental Social Psychology*, *49*(4), 771–775. https://doi.org/10.1016/J.JESP.2013.02.003

Sutherland, C. A. M., Liu, X., Zhang, L., Chu, Y., Oldmeadow, J. A., & Young, A. W. (2018). Facial first impressions across culture: Data-driven modeling of chinese and british perceivers' unconstrained facial impressions. *Personality and Social Psychology Bulletin*, *44*(4), 521–537. https://doi.org/10.1177/0146167217744194

Sutherland, C. A., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*(1), 105–118. https://doi.org/10.1016/J.COGNITION.2012.12.001

Sutherland, C. A., Oldmeadow, J. A., & Young, A. W. (2016). Integrating social and facial models of person perception: Converging and diverging dimensions. *Cognition*, *157*, 257–267. https://doi.org/10.1016/j.cognition.2016.09.006

Sutherland, C. A., Rhodes, G., Burton, N. S., & Young, A. W. (2020). Do facial first impressions reflect a shared social reality? *British Journal of Psychology*, *111*(2), 215–232. https://doi.org/10.1111/bjop.12390

Sutherland, C. A., Young, A. W., Mootz, C. A., & Oldmeadow, J. A. (2015). Face gender and stereotypicality influence facial trait evaluation: Counter-stereotypical female faces are negatively evaluated. *British Journal of Psychology*, *106*(2), 186–208. https://doi.org/10.1111/bjop.12085

Sutherland, C. A., Young, A. W., & Rhodes, G. (2017). Facial first impressions from another angle: How social judgements are influenced by changeable and invariant facial properties. *British Journal of Psychology*, *108*(2), 397–415. https://doi.org/10.1111/bjop.12206

Swencionis, J. K., Dupree, C. H., & Fiske, S. T. (2017). Warmth-competence tradeoffs in impression management across race and social-class divides. *Journal of Social Issues*, *73*(1), 175–191. https://doi.org/10.1111/JOSI.12210

Tan, Y., Rérolle, S., Lalitharatne, T. D., van Zalk, N., Jack, R. E., & Nanayakkara, T. (2022). Simulating dynamic facial expressions of pain from visuo-haptic interactions with a robotic patient. *Scientific Reports*, *12*(1), 1–11. https://doi.org/10.1038/s41598-022-08115-1

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, *23*(3), 457–482. https://doi.org/10.1016/0010-0285(91)90016-H

Terburg, D., Hooiveld, N., Aarts, H., Kenemans, J. L., & van Honk, J. (2011). Eye tracking unconscious face-to-face confrontations: Dominance motives prolong gaze to masked angry faces. *Psychological Science*, *22*(3), 314–319. https://doi.org/10.1177/0956797611398492

Thoresen, J. C., Vuong, Q. C., & Atkinson, A. P. (2012). First impressions: Gait cues drive reliable trait judgements. *Cognition*, *124*(3), 261–271. https://doi.org/10.1016/J.COGNITION.2012.05.018

Thorstenson, C. A. (2018). The social psychophysics of human face color: Review and recommendations. *Social Cognition*, *36*(2), 247–273. https://doi.org/10.1521/soco.2018.36.2.247

Thorstenson, C. A., Elliot, A. J., Pazda, A. D., Perrett, D. I., & Xiao, D. (2018). Emotion-color associations in the context of the face. *Emotion*, *18*(7), 1032–1042. https://doi.org/10.1037/EMO0000358

Thorstenson, C. A., & Pazda, A. D. (2021). Facial coloration influences social approach-avoidance through social perception. *Cognition and Emotion*, *35*(5), 970–985. https://doi.org/10.1080/02699931.2021.1914554

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: Complementarity in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology*, *84*(3), 558–568. https://doi.org/10.1037/0022-3514.84.3.558

Tinbergen, N., & Tinbergen, N. (1948). Social releasers and the experimental method required for their study. *The Wilson bulletin.*, *60*(1), 6–51. https://www.biodiversitylibrary.org/part/208759

Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*, *1124*, 208–224. https://doi.org/10.1196/annals.1440.012

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

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*(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, *27*(6), 813–833. http://haxbylab.dartmouth.edu/publications/TPO09.pdf

Todorov, A., & Porter, J. M. (2014). Misleading First Impressions. *Psychological Science*, *25*(7), 1404–1417. https://doi.org/10.1177/0956797614532474

Toscano, H., Schubert, T. W., & Giessner, S. R. (2018). Eye gaze and head posture jointly influence judgments of dominance, physical strength, and anger. *Journal of Nonverbal Behavior*, *42*(3), 285–309. https://doi.org/10.1007/S10919-018-0276-5/TABLES/2

Tracy, J. L., & Robins, R. W. (2008). The nonverbal expression of pride: evidence for cross-cultural recognition. *Journal of personality and social psychology*, *94*(3), 516–530. https://doi.org/10.1037/0022-3514.94.3.516

Traum, D. R., Swartout, W., Marsella, S., & Gratch, J. (2005). Virtual humans for non-team interaction training. *AAMAS Workshop on Creating Bonds with Embodied Conversational Agents*.

Trope, Y., & Higgins, E. T. (1993). The what, when, and how of dispositional inference: New answers and new questions. *Personality and Social Psychology Bulletin*, *19*(5), 493–500. https://doi.org/10.1177/0146167293195002

Tsankova, E., & Kappas, A. (2016). Facial skin smoothness as an indicator of perceived trustworthiness and related traits. *Perception*, *45*(4), 400–408. https://doi.org/10.1177/0301006615616748

Valentine, K. A., Li, N. P., Meltzer, A. L., & Tsai, M. H. (2020). Mate Preferences for Warmth-Trustworthiness Predict Romantic Attraction in the Early Stages of Mate Selection and Satisfaction in Ongoing Relationships. *Personality and Social Psychology Bulletin*, *46*(2), 298–311. https://doi.org/10.1177/0146167219855048

Valentine, K. A., Li, N. P., Meltzer, A. L., Tsai, M. H., Yong, J. C., Tov, W., Sng, O., Fletcher, G. J. O., Valentine, K. A., Jiang, Y. F., & Balliet, D. (2013). Mate preferences do predict attraction and choices in the early stages of mate selection. *Journal of Personality and Social Psychology*, *105*(5), 757–776. https://doi.org/10.1037/a0033777

Valentine, K. A., Li, N. P., Penke, L., & Perrett, D. I. (2014). Judging a man by the width of his face: The role of facial ratios and dominance in mate choice at speed-dating events. *Psychological Science*, *25*(3), 806–811. https://doi.org/10.1177/0956797613511823

Varghese, L., Irene Huntoon Lindeman, M., & Finkelstein, L. (2018). Dodging the double bind: the role of warmth and competence on the relationship between interview communication styles and perceptions of women's hirability. *European Journal of Work and Organizational Psychology*, *27*(4), 1464–0643. https://doi.org/10.1080/1359432X.2018.1463989

Varnum, M. E. W., & Denson, T. (2013). What are lay theories of social class? *PLOS ONE*, *8*(7), e70589. https://doi.org/10.1371/JOURNAL.PONE.0070589

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 of the United States of America*, *111*(32), E3353–61. https://doi.org/10.1073/pnas.1409860111

Walker, M., & Vetter, T. (2009). Portraits made to measure: Manipulating social judgments about individuals with a statistical face model. *Journal of Vision*, *9*(11), 1–13. https://doi.org/10.1167/9.11.1

Walker, M., & Vetter, T. (2016). Changing the personality of a face: Perceived big two and big five personality factors modeled in real photographs. *Journal of Personality and Social Psychology*, *110*(4), 609–624. https://doi.org/10.1037/pspp0000064

Wang, H., Han, C., Hahn, A. C., Fasolt, V., Morrison, D. K., Holzleitner, I. J., DeBruine, L. M., & Jones, B. C. (2019). A data-driven study of Chinese participants' social judgments of Chinese faces (P. Sorokowski, Ed.). *PLOS ONE*, *14*(1), e0210315. https://doi.org/10.1371/journal.pone.0210315

Wang, J., Lin, W., Fang, X., & Mo, L. (2020). The influence of emotional visual context on the judgment of face trustworthiness. *Psychology Research and Behavior Management*, *13*, 963. https://doi.org/10.2147/PRBM.S269543

Watkins, C. D., Debruine, L. M., Feinberg, D. R., & Jones, B. C. (2013). A sex difference in the context-sensitivity of dominance perceptions. *Evolution and Human Behavior*, *34*(5), 366–372. https://doi.org/10.1016/J.EVOLHUMBEHAV.2013.06.004

Watkins, C. D., Jones, B. C., & DeBruine, L. M. (2010). Individual differences in dominance perception: Dominant men are less sensitive to facial cues of male dominance. *Personality and Individual Differences*, *49*(8), 967–971. https://doi.org/10.1016/J.PAID.2010.08.006

Watt, M. C., Maitland, R. A., & Gallagher, C. E. (2017). A case of the "Heeby Jeebies": An examination of intuitive judgements of "Creepiness". *Canadian Journal of Behavioural Science-Revue Canadienne Des Sciences Du Comportement*, *49*(1), 58–69. https://doi.org/10.1037/cbs0000066

Wayne Leach, C., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, *93*(2), 234–249. https://doi.org/10.1037/0022-3514.93.2.234

Wen, F., Zuo, B., Ma, S., Xu, Y., Coley, J. D., & Wang, Y. (2020). Do we see masculine faces as competent and feminine faces as warm? Effects of sexual dimorphism on facial perception. *Evolutionary Psychology*, *18*(4), 147470492098064. https://doi.org/10.1177/1474704920980642

Whitty, G. (2001). Education, social class and social exclusion. *Journal of Education Policy*, *16*(4), 287–295. https://doi.org/10.1080/02680930110054308

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, *37*(3), 395–412. https://doi.org/10.1037/0022-3514.37.3.395

Williams, L. M., Senior, C., David, A. S., Loughland, C. M., & Gordon, E. (2001). In search of the "Duchenne smile": Evidence from eye movements. *Journal of Psychophysiology*, *15*(2), 122–127. https://doi.org/10.1027//0269-8803.15.2.122

Willis, J., & Todorov, A. (2006). First impressions making up your mind after a 100-Ms exposure to a face. *Psychological Science*, *17*(7), 592–598. https://doi.org/https://doi.org/10.1111/j.1467-9280.2006.01750.x

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, *26*(8), 1325–1331. https://doi.org/10.1177/0956797615590992

Witkower, Z., & Tracy, J. L. (2019). A facial-action imposter: How head tilt influences perceptions of dominance from a neutral face. *Psychological Science*, *30*(6), 893–906. https://doi.org/10.1177/0956797619838762

Wojciszke, B., Abele, A. E., & Baryla, W. (2009). Two dimensions of interpersonal attitudes: Liking depends on communion, respect depends on agency. *European Journal of Social Psychology*, *39*(6), 973–990. https://doi.org/10.1002/EJSP.595

Wolff, S. E., & Puts, D. A. (2010). Vocal masculinity is a robust dominance signal in men. *Behavioral Ecology and Sociobiology*, *64*(10), 1673–1683. https://doi.org/10.1007/S00265-010-0981-5/TABLES/4

Wolffhechel, K., Fagertun, J., Jacobsen, U. P., Majewski, W., Hemmingsen, A. S., Larsen, C. L., Lorentzen, S. K., & Jarmer, H. (2014). Interpretation of appearance: The effect of facial features on first impressions and personality. *PLOS ONE*, *9*(9), e107721. https://doi.org/10.1371/journal.pone.0107721

Xie, S. Y., Flake, J. K., & Hehman, E. (2019). Perceiver and target characteristics contribute to impression formation differently across race and gender. *Journal of Personality and Social Psychology*, *117*(2), 364–385. https://doi.org/10.1037/pspi0000160

Xie, S. Y., Flake, J. K., Stolier, R. M., Freeman, J. B., & Hehman, E. (2021). Facial impressions are predicted by the structure of group stereotypes. *Psychological Science*, *32*(12), 1979–1993. https://doi.org/10.1177/09567976211024259

Yu, H., Garrod, O. G., & Schyns, P. G. (2012). Perception-driven facial expression synthesis. *Computers and Graphics (Pergamon)*, *36*(3), 152–162. https://doi.org/10.1016/j.cag.2011.12.002

Zambelli, A. E. (2016). A data-driven approach to estimating the number of clusters in hierarchical clustering. *F1000Research*, *5*. https://doi.org/10.12688/F1000RESEARCH.10103.1

Zebrowitz, L. A. (2017). First Impressions From Faces. *Current Directions in Psychological Science*, *26*(3), 237–242. https://doi.org/10.1177/0963721416683996

Zebrowitz, L. A. (2018). *Reading faces: Window to the soul?* (1st Editio). Routledge. https://doi.org/10.4324/9780429493188

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., & Franklin, R. G. (2014). The attractiveness halo effect and the babyface stereotype in older and younger adults: Similarities, own-age accentuation, and older adult positivity effects. *Experimental Aging Research*, *40*(3), 375–393. https://doi.org/10.1080/0361073X.2014.897151

Zebrowitz, L. A., Franklin, R. G., Hillman, S., & Boc, H. (2013). Older and younger adults' first impressions from faces: Similar in agreement but different in positivity. *Psychology and Aging*, *28*(1), 202–212. https://doi.org/10.1037/A0030927

Zebrowitz, L. A., Franklin Jr., R. G., Boshyan, J., Luevano, V., Agrigoroaei, S., Milosavljevic, B., & Lachman, M. E. (2014). Older and younger adults' accuracy in discerning health and competence in older and younger faces. *Psychology and aging*, *29*(3). https://doi.org/10.1037/A0036255

Zebrowitz, L. A., Kikuchi, M., & Fellous, J. M. (2007). Are effects of emotion expression on Trait Impressions mediated by babyfaceness? evidence from connectionist modeling. *Personality and Social Psychology Bulletin*, *33*(5), 648–662. https://doi.org/10.1177/0146167206297399

Zebrowitz, L. A., Kikuchi, M., & Fellous, J.-M. (2010). Facial resemblance to emotions: group differences, impression effects, and race stereotypes. *Journal of personality and social psychology*, *98*(2), 175–189. https://doi.org/10.1037/a0017990

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., Wadlinger, H. A., Luevano, V. X., White, B. M., Xing, C., & Zhang, Y. (2011). Animal analogies in first impressions of faces. *Social Cognition*, *29*(4), 486–496. https://doi.org/10.1521/soco.2011.29.4.486

Zhan, J., Garrod, O. G., van Rijsbergen, N., & Schyns, P. G. (2019). Modelling face memory reveals task-generalizable representations. *Nature Human Behaviour*, *3*(8), 817–826. https://doi.org/10.1038/s41562-019-0625-3

Zhan, J., Liu, M., Garrod, O. G., Daube, C., Ince, R. A., Jack, R. E., & Schyns, P. G. (2021). Modeling individual preferences reveals that face beauty is not universally perceived across cultures. *Current Biology*, *31*(10), 2243–2252.e6. https://doi.org/10.1016/J.CUB.2021.03.013

Zhang, D., Lin, H., & Perrett, D. I. (2020). Apparent emotional expression explains the effects of head posture on perceived trustworthiness and dominance, but a measure of facial width does not. *Perception*, *49*(4), 422–438. https://doi.org/10.1177/0301006620909286

Zhang, Y., & Ji, Q. (2005). Active and dynamic information fusion for facial expression understanding from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(5), 699–714. https://doi.org/10.1109/TPAMI.2005.93

Zhong, V. J., Mürset, N., Jäger, J., & Schmiedel, T. (2022). Exploring variables that affect robot likeability. *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, 1140–1145. https://doi.org/10.5555/3523760