LINKING 3D FACE SHAPE TO SOCIAL PERCEPTION

Iris J. Holzleitner

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2015

LINKING 3D FACE SHAPE TO SOCIAL PERCEPTION

Iris J. Holzleitner

University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of
PhD
at the
University of St Andrews

Date of Submission
21 September 2015

Declarations

Candidate's declaration

I, Iris J. Holzleitner, hereby certify that this thesis, which is approximately 40,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in October 2011 and as a candidate for the degree of PhD in July 2012; the higher study for which this is a record was carried out in the University of St Andrews between 2011 and 2015.

Date ______________________ Signature of candidate ______________________

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date ______________________ Signature of supervisor ______________________

Permission for publication

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED AND ELECTRONIC COPY
Embargo on all of printed and electronic copy for a period of 2 years (maximum five) on the following ground(s):

- Publication would preclude future publication

Supporting statement for print and electronic embargo request
Substantial parts of the empirical work in this thesis have not yet been published in peer-reviewed journals. Publication of articles can have long lag periods, thus two years may be required to publish all results of this thesis.

Date: September 21, 2015
Signature of candidate ________________ Signature of supervisor ________________

Collaboration statement

Throughout the experimental chapters in this thesis, I have used the pronoun "we" in addition to "I". This work is my own under the support of my supervisor in terms of hypotheses, experimental design, analyses and conclusions; however, the Perception Lab is an inherently collaborative environment. The plural pronoun reflects the fact that if/when published, the following experiments would carry multiple authorship and is used in keeping with intellectual honesty.

Acknowledgments

While PhD-ing can be a rough business at times, a great many people have made the last four years a hugely enjoyable experience. I am particularly indebted to the following people.

To Dave Perrett, for being a fabulous mentor, never getting tired of discussing ideas—promising or not so much—and challenging me. I could not have wished for a better supervisor. Thank you for providing us all with an immensely stimulating and fun environment to work in (as well as an unlimited supply of coffee, and treats from the garden).

To the two generations of Perception Lab members I had the pleasure of sharing days (and nights) in the lab with, Amanda, Carmen, Dan, Milena, and especially Audrey, Carlot(t)a, Jing, Martha and Sean(n). I'll miss the banter (and bickering). Special thanks to Willypete for taking me out for wee breaks.

To Lesley Ferrier, for being the calm rock on which we all rely (and who yet manages to never lose her good spirits), Dengke Xiao, for his computing wizardry, and Anne Perrett for her eagle-eyed proof-reading of manuscripts. My apologies for all the "howevers".

To Bernie Tiddeman, for his ceaseless patience and swift help whenever I ran into a problem with MorphAnalyser, and Dave Hunter, for his tech and stats support when I first joined the lab.

To my St Andrews family, Cathy and Lizzie for never failing to cheer me up and helping me to keep sane (mostly), and Malcolm, Naomi, Jamie and Becky for being great friends and mentors.

To my sister Gerda, for her generous loan.

To my mum, for being my number one fan and supporting me all the way.

And, last but not least, to Ross, for freeing his desk for me, reminding me that there are happy places whenever my pessimism kicked in and I lost sight of them, and for exploring Scotland's remotest corners with me. Here's to exploring some more; I can't wait.

Abstract

Advances in computer graphic and statistical methods have made it possible to visualise global face shape correlates of social judgments. The current thesis used a data-driven approach to investigate face shape correlates and perception of two traits, masculinity and strength, both of which are important in mate choice and social perception more generally. The studies presented defined the influences of body physique (height, body mass index, body fat and muscle mass) on facial shape, and their effects on the perception of masculinity, attractiveness and strength.

Study 1 investigated the face shape correlates of actual and perceived masculinity. I found that perceived masculinity is not only driven by sexually dimorphic shape, but also by cues to body height and weight. Men with taller and heavier bodies were perceived to have more masculine-looking faces.

Study 2 investigated women's perception of male attractiveness as a function of masculine face shape. As previously assumed but not explicitly tested, I found that masculinity preferences followed a quadratic relationship: attractiveness increased with increasing masculinity levels, but dropped off at higher levels of masculinity. In addition, I showed that the relative costs and benefits of high and low masculinity are affected by individual differences in own condition, perceived financial harshness and pathogen disgust.

In Study 3, I found that perception of strength from faces is driven by facial cues to body physique; individuals with higher body bulk were perceived to be stronger. In men, it proved possible to further dissociate facial cues to muscle and fat mass which both contributed to strength perception.

The thesis demonstrates that facial cues used in the evaluation of masculinity and strength are linked to bodily characteristics associated with sex differences and actual strength, namely height, weight, muscularity and adiposity. My findings therefore support the hypothesis that perceptions have an adaptive origin.

Contents

[1] Part of the work in this section has been published in *Perception*. The reference for this work is as follows: Holzleitner, I. J., Hunter, D. W., Tiddeman, B. P., Seck, A., Re, D. E., & Perrett, D. I. (2014). Men's facial masculinity: when (body) size matters. *Perception, 43*, 1191-1202.

[2] Part of the work in this section is currently under review: Holzleitner, I. J., & Perrett, D. I. (under review). Perception of strength from 3D faces is linked to facial cues to physique. *Evolution and Human Behavior*.

List of Figures

List of Tables

Part I.

General Introduction

1. The relevance of faces in social interactions

To "judge a book by its cover" is considered ill-advised and socially undesirable. Yet, numerous studies show that facial appearance affects how we perceive others, especially on—but not limited to—first encounters. Faces are central to social interactions. As Hassin and Trope (2000) noted, they are available in almost every social situation and provide relatively stable information regarding a range of socially relevant categories such as sex and age (Bruce & Young, 1986), but can also reveal emotions and intentions. The importance of faces also is supported by neuroscientific evidence showing that some brain areas are specialized for processing faces (e.g., Kanwisher & Barton, 2011). Inferences from faces go far beyond basic judgments such as sex or age: One study found that 75% of a sample of more than 500 adults explicitly embraced the notion that faces reveal "true" personality (Hassin & Trope, 2000). The idea that facial appearance might be linked to behavioural traits or personality is termed physiognomy, and can be traced back to antiquity. Physiognomy received renewed interest in the 18th and 19th century, but was soon dismissed as pseudoscience. Berry and Wero (1993) identified fear of association with physiognomists as one of the reasons why psychologists were initially reluctant to investigate the link of physical appearance and impression formation. Yet, empirical evidence gathered over the last decades suggests that some initial perceptions not only show high inter-rater agreement, but also might have a "kernel of truth" to them (Berry & Wero, 1993).

This thesis is based on the premise that understanding *how* inter-personal impressions are formed from faces helps to understand *why* they are formed, and one interest of this thesis lies in linking these *whys* to evolutionary reasoning. This introductory chapter will first briefly review the relevance of faces in social interactions, theoretical and methodological approaches to social face perception, and then introduce the scope of the current thesis.

1.1. Appearance-based judgments affect social interactions

One of the most researched aspects of facial appearance is facial attractiveness. The power of physical attractiveness in impression formation and social interactions has been well documented. Even infants distinguish between attractive and unattractive faces (Langlois et al., 1987), and from infancy throughout adult life more attractive people are perceived and treated preferentially (Dion, Walster & Berscheid, 1972; Feingold, 1992; Langlois et al., 2000) in academic (e.g., Ritts, Patterson & Tubbs, 1992), occupational (Cash, Gillen & Burns, 1977; Dipboye, Arvey & Terpstra, 1977; Dipboye, Fromkin & Wiback, 1975; Hamermesh & Biddle, 1994) and mating contexts (Jokela, 2009; Pflüger, Oberzaucher, Katina, Holzleitner & Grammer, 2012; Walster, Aronson, Abrahams & Rottmann, 1966) and even in court (e.g., Sigall & Ostrove, 1975).

Facial attractiveness is not the only appearance-based judgment that has been shown to have real-world outcomes. Like judgments of facial attractiveness, social attributions, too, have been found to show high inter-rater agreement (e.g., Engell, Haxby & Todorov, 2007; Zebrowitz McArthur & Berry, 1987) and to impact strongly on various social outcomes. Facial perceptions of competence predict the selection of political (e.g., Ballew & Todorov, 2007; Olivola & Todorov, 2010a; Rule et al., 2010; Todorov, Mandisodza, Goren & Hall, 2005) as well as corporate leaders: the CEOs of more successful companies look like better leaders (Rule & Ambady, 2008, 2009), and more competent-looking CEOs receive higher salaries (Graham, Harvey & Puri, 2010). Zebrowitz and Montepare (2005) suggested that perceptions of competence reflect difference in babyfaceness; earlier studies by Zebrowitz and colleagues found evidence suggesting that babyfaceness, too, affects hiring (Collins & Zebrowitz, 1995; Zebrowitz, Tenenbaum & Goldstein, 1991) and sentencing (Berry & Zebrowitz McArthur, 1988; Zebrowitz & McDonald, 1991) decisions.

Trustworthiness is another social attribute that has powerful effects on social interactions including sentencing: Wilson and Rule (2015), for example, showed that facial trustworthiness predicts the likelihood of death sentences for convicted murderers, and perceptions of trustworthiness were found to predict electoral outcomes in Japan (Rule et al., 2010). On an inter-personal level, trustworthiness judgments likely affect who we choose to cooperate with—experimental studies show that untrustworthy-looking individuals are less likely trusted in economic games (e.g., Rezlescu, Duchaine, Olivola & Chater, 2012). Even children as young as five years old appear to base their cooperative behaviour on facial impressions of trustworthiness (Ewing, Caulfield, Read & Rhodes, 2014).

Hassin and Trope (2000) found that facially inferred personality can change the interpretation of verbal information, and that facial appearance has consistent and uncontrollable effects on decisions (even when asked to ignore facial photographs of hypothetical job seekers, participants failed to do so). While their study suggests that we do read from faces, they also found evidence that we read into faces: perceptions of faces can be altered when information about personality is available (Hassin & Trope, 2000)[1]. Nonetheless, Rudoy and Paller (2009) showed that perceptual information from facial images was more influential than verbal cues

[1] Compare findings by Little, Burt and Perrett (2006) who showed that faces perceived to reflect desired personality traits were rated as more attractive.

to trustworthiness when judgments were made under time constraints. Olivola and Todorov (2010b) found that facial appearance prevails over other social cues in driving judgments, even if those other cues might be more valid. Similarly, Rezlescu et al. (2012) showed that the effect of a trustworthy facial appearance in a trust game was greatly reduced but remained significant when behavioural information was available.

In summary, appearance-based impression formation is of significant importance in social interactions.

1.2. Faces elicit snap judgments

Willis and Todorov (2006) found that an exposure time of 100 ms was enough for observers to form judgements of attractiveness, likability, trustworthiness, competence and aggressiveness that correlated with ratings by a different set of raters without time constraints. Increasing exposure time to up to 1 s did not significantly change judgments, but increased confidence in judgments and lead to more differentiated person impression—the correlation between the five different judgments decreased with longer exposure, and so did the effect of attractiveness on trait judgments (see also Todorov, Pakrashi & Oosterhof, 2009). Bar, Neta and Linz (2006) found that observers formed relatively stable impressions of how threatening a (neutral) unknown face looks after exposure times as little as 39 ms. Rapidity of judgments has also been demonstrated for ratings of competence (Ballew & Todorov, 2007) and extraversion (Borkenau, Brecke, Möttig & Paelecke, 2009), amongst others, suggesting that certain judgments are made spontaneously and automatically.

1.3. Faces elicit (somewhat) accurate judgments

Research on the accuracy of social perceptions has often focussed on the "Big Five" personality characteristics (openness, conscientiousness, extraversion, agreeableness, neuroticism), as well as dimensional traits derived from circumplex models (e.g., Wiggins, 1979) such as being submissive versus dominant. By relating trait judgments to self-report measures of personality, small to moderate correlations have been found between observer- and self-rated measures of personality for some but not other traits. For example, in studies reviewed by Zebrowitz and Collins (1997) correlations of other-rated dominance and self-report questionnaire measures ranged from .11 for women up to .53 for men. Extraversion was found to be perceived with some accuracy by Penton-Voak, Pound, Little and Perrett (2006) and Borkenau et al. (2009)[2]. Berry and Wero (1993) found that observers were able to predict social dominance, interpersonal

[2]Note, however, that these findings were based on unstandardized stimulus pictures—accuracy for judgments of extraversion was mediated by facial expressions

warmth and honesty with some accuracy. Limited accuracy was also found for judgments of intelligence (Zebrowitz, Hall, Murphy & Rhodes, 2002) and perceived trustworthiness (Porter, England, Juodis, ten Brinke & Wilson, 2008).

Studies on the accuracy of judgments might have over-estimated true accuracy. Todorov, Olivola, Dotsch and Mende-Siedlecki (2015) identified the following shortcomings of previous studies on judgment accuracy. First, a failure to control for gender, ethnicity and age which may provide obvious indicators of the traits being inferred; second, studies have compared accuracy of face judgments against chance when it should be compared against other sources of information (such as base rates of traits in the population). Third, it is usually assumed that facial images adequately capture facial appearance—yet, there can be considerable variation between images of the same individual (e.g., Jenkins, White, Van Montfort & Burton, 2011; Todorov & Porter, 2014).[3] Nonetheless, evidence does suggest that judgments are not completely arbitary.

In summary, judgments such as trustworthiness are made automatically and rapidly, seem to prevail even when additional (conflicting) behavioural information is available and have a strong impact on social decisions and interactions—but if these judgments are of limited (and sometimes indeed very poor) accuracy, why do we make them?

[3] Note that this variation might be of less concern when capturing participants with neutral facial expressions, standardising head posture and excluding non-face cues such as clothing and hair, which are standard procedures in many face perception studies.

2. Reading faces: understanding social face perception

The following section will review research trying to answer why we form quick and sometimes persistent impressions if they are not necessarily accurate. The main premise of the reviewed research is that faces may provide information regarding future social interactions which is adaptive to attend to; some facial cues might have evolutionary significance. The following section will first briefly discuss theoretical frameworks. The ecological theory of face perception emphasises that perception impacts on behaviour and social interactions; it intersects with evolutionary psychology theories that focus on the importance of facial cues in the context of sexual selection and mate choice (Zebrowitz & Montepare, 2008). The second part of the following section will then introduce different methodological frameworks that have been deployed to investigate social face perception.

2.1. Theoretical frameworks

2.1.1. Perceiving is for doing: adaptive perceptions and overgeneralizations

The ecological theory of social perception suggests that observers are attuned to cues that have adaptive significance (Zebrowitz, 2011; Zebrowitz & Collins, 1997; Zebrowitz, Fellous, Mignault & Andreoletti, 2003; Zebrowitz McArthur & Baron, 1983). From an evolutionary point of view, it would have been important to detect with accuracy those attributes that affect the outcomes of potential social interaction. For example, individuals that pose a threat should be avoided. If the costs of missing a cue to threat are higher than those of erroneously ascribing threat to someone that happens to bear facial traits resembling cues to threat, a bias in social perception, i.e. overgeneralization, can arise (Zebrowitz & Collins, 1997; Zebrowitz & Montepare, 2008). Zebrowitz (1996) proposed that facial traits resembling emotions might be overgeneralized in trait judgments.[1] In line with this reasoning, Said, Sebe and Todorov

[1] Previously, Secord (1958) had proposed that emotional expressions are perceived to extend in time: current expression can lead to mistaken inferences about stable traits. Knutson (1996), for example, found that people with angry expressions were perceived high in dominance, while people with sad or fearful expression were perceived low in dominance.

(2009), for example, found that (neutral) faces that showed structural resemblance to angry facial expressions were perceived to bear negative (threatening) traits (see also Montepare & Dobish, 2003; Neth & Martinez, 2009). Emotion-related trait inferences also appear to interact with facial cues to sex. Hess, Adams Jr., Grammer and Kleck (2009) hypothesized that because expressive markers of anger overlap with markers of sex (e.g., lowered eyebrows), androgynous angry faces would be more likely perceived as male than female; indeed, they found that angry faces were more likely perceived to be men (while happy/fearful faces were more likely to be perceived as women) (see also Zebrowitz, Kikuchi & Fellous, 2010).

One of the most compelling studies in support of a functional basis of face perception (and tying in with reasoning on emotion overgeneralization) was conducted by Oosterhof and Todorov (2008). Research on social perception had suggested that faces are evaluated along two dimensions (trustworthiness or valence, and dominance). Oosterhof and Todorov (2008) showed that these perceptions might represent an overgeneralization of cues indicative of harmful intentions and the ability to cause harm, respectively. Whereas the valence dimension comprised facial cues to expressions signalling approach and avoidance behaviours, the dominance dimension appeared to comprise cues to physical strength: The valence but not the dominance dimension was related to judgments on a scale ranging from angry to happy, whereas ratings of facial maturity and femininity/masculinity where more strongly related to the dominance than to the valence dimension (see also Oosterhof & Todorov, 2009).[2]

2.1.2. Mate preferences: Darwinian aesthetics

Owing to its importance in social interactions, facial attractiveness has been heavily researched. Based on findings that from an early age and across diverse cultures there is noteworthy agreement in ratings of facial attractiveness (Langlois et al., 2000; Rhodes, 2006), a lot of research has investigated attractiveness within an evolutionary framework. It has been hypothesized that "[w]hen members of a species discriminate between potential mates with regard to their physical appearance, as humans do, . . . the discrimination reflects special-purpose adaptations responsive to cues that had mate value in evolutionary history" (Thornhill & Gangestad, 1999, p. 452). Put differently, mating decisions, advised by specific preferences and affecting reproductive success (Jokela, 2009; Pflüger et al., 2012; Rhodes, Simmons & Peters, 2005), are likely to be shaped by processes of sexual selection rather than being arbitrary concepts exclusively shaped by cultural norms (Grammer, Keki, Striebel, Atzmüller & Fink, 2003; Rhodes, 2006; Thornhill & Gangestad, 1999).

[2] A recent study using less controlled stimulus images suggests that a third dimension which the authors termed "youthful-attractiveness" might be used to evaluate faces (Sutherland et al., 2013).

2.1.2.1. Sexual selection

Being startled by traits that seemed detrimental to survival (such as the much-cited peacock's tail), Darwin developed the idea of sexual selection. Its agents are sexual rivals and mates (Ghiselin, 1974, as cited in Andersson, 1994)—sexual selection, as conceptualised by Darwin, is driven by males competing against each other for females (intrasexual selection), and by females choosing males (intersexual selection).

Bateman (1948) sought to provide an explanation for intrasexual competition, and found that the variance in reproductive success differed more between male than between female *Drosophila*. Whereas male reproductive success increases with the number of matings achieved, female reproductive success does not increase after the first mating. *Drosophila* female reproductive success is limited by the number of eggs that can be produced, but male reproductive success is only limited by the ability to fertilise eggs, leaving males with a higher reproductive potential than females and a fierce competition for mates. Trivers (1972) defined a more general principle governing sexual selection: the relative parental investment in offspring by the two sexes. He defined parental investment as "any investment by the parent in an individual offspring that increases the offspring's chance of surviving (and hence reproductive success) at the cost of the parent's ability to invest in other offspring" (Trivers, 1972, p. 139). What governs sexual selection, thus, is the relative investment of the sexes[3]: The sex to invest more should be more critical in mate choice; the sex to invest less should be competitive about sexual access to members of the higher-investing sex. In humans, it is women who have a higher obligatory investment in offspring (cost-intensive gametes, gestation, lactation, etc.), and a limited fertile phase.[4]

Traits whose expression is related to differences in reproductive success, and caused by competition over mates (Andersson, 1994) are called sexually selected traits. *Intra*sexually selected traits are sexually dimorphic traits that benefit males in aggressive encounters with same-sex individuals, such as increased body height. *Inter*sexually selected traits can benefit their bearers via two different routes: first, through Fisherian runaway selection (Fisher, 1930) for "sexy sons" (and daughters, Cornwell & Perrett, 2008), which leads to the evolution of ornaments—traits whose value is initially arbitrary, but increases offspring's reproductive success by rendering them more desirable to potential partners. And second, opting mate choice for honest signals, leading to traits that indicate good genes or non-heritable benefits like parental care, providing abilities, protective behavior etc. (e.g., Hamilton & Zuk, 1982;

[3] Although it has been suggested that operational sex ratios and potential reproductive rates might be other important determinants (e.g., Eens & Pinxten, 2000).

[4] Note that this is not the case in all species (sex-role reversal, e.g., Andersson, 1994), and even *within* species it has been observed that populations may switch between exhibiting conventional and reversed sex-roles depending on fluctuations in ecological or physical factors (e.g., Eens & Pinxten, 2000), and this may also apply to humans (Brown, Laland & Mulder, 2009).

Trivers, 1972; Williams, 1966; Zahavi & Zahavi, 1997). Often, ornaments and indicators are not easily discriminable. According to Miller, "[a]ny particular trait that evolved through sexual selection was probably influenced by some combination of runaway processes, pressures to advertise fitness, and psychological preferences." (Miller, 2001, p. 159–161). Indeed, Kokko, Brooks, Jennions and Morley (2003) argue the dichotomy of "Fisherian runaway" and Zahavian "good genes" models to be misrepresentative, as both preferences and traits become correlated when they have a heritable genetic basis.

Several such sexually selected traits have been put forward, such as symmetry, (e.g., Grammer & Thornhill, 1994; B. C. Jones, Little & Perrett, 2003; Moller & Thornhill, 1998; Perrett et al., 1999; Thornhill & Gangestad, 1993), averageness (e.g., Alley & Cunningham, 1991; Grammer et al., 2003; Grammer & Thornhill, 1994; Langlois & Roggman, 1990; Rhodes & Tremewan, 1996; Thornhill & Gangestad, 1993), skin texture (e.g., Fink, Grammer & Matts, 2006; B. C. Jones, Little, Burt & Perrett, 2004; S. C. Roberts et al., 2005), scent (e.g., Gangestad & Thornhill, 1998; Grammer, Fink & Neave, 2005; Rikowski & Grammer, 1999; Thornhill & Gangestad, 1999), and movement quality, (e.g., Fink, Weege, Neave, Pham & Shackelford, 2015; Grammer et al., 2003). Another trait, and of main interest to the current work, is (facial) sexual dimorphism, which will be briefly reviewed in the following section.

2.1.2.2. Facial sexual dimorphism

Proximately, dimorphism in secondary sexual traits is caused by sex-specific ratios of two classes of steroid hormones, androgens and oestrogens. Both sexes have both hormones, but the ratio at puberty is sex-specific (Thornhill & Moller, 1997). Sex hormones affect both morphology and behaviour via multiple routes. Prenatally, they have organising effects on body and brain anatomy (e.g., Mazur & Booth, 1998; McEwen, 1988). During puberty, sexual dimorphism increases drastically under their influence, and fluctuating levels of sex hormones continue to affect human behaviour post-pubertally (see Section 7). Hence, pre- and post-natal hormone levels, behaviour and (face-) morphological features are mutually linked.

In women, typically female (facial) features have been linked to higher attractiveness in numerous studies (Cunningham, 1986; Fraccaro et al., 2010; Johnston & Franklin, 1993; Johnston, Hagel, Franklin, Fink & Grammer, 2001; D. Jones & Hill, 1993; Law Smith et al., 2006; Perrett et al., 1998; Perrett, May & Yoshikawa, 1994; Rhodes, 2006; Rhodes, Hickford & Jeffery, 2000). Female femininity is linked (and sometimes even treated synonymously, D. Jones & Hill, 1993) to neotenous traits:[5] big eyes, full lips, a small nose, and reduced vertical dimensions (Cunningham, 1986; D. Jones et al., 1995; D. Jones & Hill, 1993). Korthase

[5] "Neoteny" in this context refers to features distinguishing older adults from young adults rather than babies.

and Trenholme (1982) found that younger female faces were rated as more attractive, and D. Jones et al. (1995) showed that men from five populations preferred female faces bearing neotenous traits as mentioned above. According to D. Jones et al. (1995), neoteny constitutes an essential part of female attractiveness, because it signals youth and thereby fertility. D. Jones (1996) argued that individual females displaying supernormal youth had an advantage in female-female competition for desirable mates. Neoteny thus would constitute a by-product of an adaptive fondness for youth, which has been tightly connected to fertility in our female ancestors. With increasing age, faces become more masculine as the oestrogen to androgen ratio lowers in women, and fertility decreases (Symons, 1995; Thornhill & Gangestad, 1999).

The most attractive female faces, however, also show markers of maturity such as high cheekbones (which are thought to develop as a result of increased pubertal oestrogen levels, Cunningham, 1986; Grammer et al., 2003), possibly because a combination of neonate and mature features might signal an optimal age for reproduction. Indeed, oestrogen levels have been found to predict ratings of facial femininity, attractiveness and health (Law Smith et al., 2006). In addition, female facial attractiveness correlates with body attractiveness (Thornhill & Gangestad, 1999) and oestrogen-dependent body traits such as breast size and waist-to-hip ratio also positively correlate with fecundity (Jasienska, Ziomkiewicz, Ellison, Lipson & Thune, 2004).

Law Smith et al. (2012) put forward another theory regarding the attractiveness of women's facial femininity. They found that both facial femininity and maternal tendencies were associated with higher oestrogen levels; facial femininity may thus serve as a cue to maternal tendencies to potential mates. However, facial femininity was not measured directly. Composites of women desiring a higher number of children were rated to look more feminine than composites of women who reported a lower ideal number of children. Stimuli provided both textural and shape cues and were unmasked, making it hard to determine whether it was indeed oestrogen-dependent face *shape* cues that led to the perception of higher femininity in women who desired more children.

In men, the attractiveness of typically male features is debated. Nonetheless, men's secondary sexual traits were long viewed as having primarily evolved under female choice: women would choose facially masculine partners as facial masculinity would act as an honest signal of health. Recently, it has been argued that men's facial masculinity is irrelevant to women's perceptions of attractiveness, and might be primarily linked to male-male competition, as evidenced by the close link of facial masculinity to perceptions of dominance and threat. This topic will be explored in more detail in Part III of this thesis. At this point, it is only noted that men's facial masculinity is interesting for two main reasons: first, it has been suggested to be linked to dominance, one of the main facial dimensions along which unfamiliar, neutral faces are assessed, and second, facial masculinity might reliably cue behavioural tendencies such as increased aggression or competitiveness via its dependence on testosterone.

2.1.3. Parallels and disagreement

Reasoning on the functional significance of cues by proponents of an ecological theory of social perception and evolutionary psychology diverge when it comes to facial attractiveness. While evolutionary psychology considers facial attractiveness to be rooted in cues to mate condition, Zebrowitz and Rhodes (2004, see also Zebrowitz, 1996 and Zebrowitz & Collins, 1997) argue that perceptions of facial attractiveness are rooted in a sickness overgeneralization effect.[6] Individuals whose faces bear traits resembling facial cues of sickness or bad genes should be avoided as they might be contagious or have poor intellectual or social skills. This argument is based on the finding that health, for example, is only accurately perceived in less attractive individuals, while very attractive individuals are not actually healthier than individuals of average attractiveness. Thus, an apparent inaccuracy in judgments (more attractive individuals being more healthy than less attractive individuals) could be explained by an overgeneralization that is rooted in an accurate, adaptive judgment—a negative response to unhealthy faces being overgeneralized to a negative response to faces that resemble those of less healthy individuals. In contrast, from an evolutionary psychology point of view unattractive people do not resemble unhealthy people, but rather are perceived as unattractive because they *are* less healthy (or of otherwise poorer condition).

Another area of potential disagreement lies in the role facial cues to maturity play in interpersonal perception. Zebrowitz (1996) suggested that it is adaptive to respond to age-related physical qualities, such as "nurturing the very young and the very old and mating with the fertile" (Zebrowitz, 1996, p. 213). An overgeneralization of age cues may lead to individuals whose facial traits resemble those of a particular age group being perceived to have psychological qualities typical of that age. In particular, studies on an overgeneralization of age-related cues have focussed on an overgeneralization of babyface-like features (Zebrowitz, 2011; Zebrowitz & Montepare, 2008). Berry and Zebrowitz McArthur (1986), for example, found that faces with childlike features were perceived as warmer and less threatening.

Zebrowitz McArthur and Berry (1987) explicitly viewed babyfaceness as a unipolar dimension, that is, not as part of an axis stretching from baby-faced to mature.[7] While Zebrowitz (1996) acknowledges that the facial and bodily characteristics that "distinguish babies from adults also differentiate women and men",[8] or, more generally, that cues to sex and age overlap,

---

[6] Zebrowitz (2011) refers to "anomalous face overgeneralization".

[7] Zebrowitz (2011) acknowledges that perceived babyfaceness and maturity lie on an axis when it comes to perceptions of dominance.

[8] Note that this definition is not identical to the definition of neotenous traits used earlier. While femininity is linked to actual and perceived youth, it would clearly not be adaptive for men to feel sexually attracted to immature women.

they concluded that it is age-related overgeneralization that drives or at least contributes to sex-typical perceptions and gender stereotypes, and not the other way round (see also Zebrowitz, 2011). While it remains to be demonstrated whether women are perceived as more trustworthy (Dzhelyova, Perrett & Jentzsch, 2012) because they look younger, or whether individuals with feminine facial features are perceived as more trustworthy because women are, on average, less likely to show aggressive/threatening behaviour, overgeneralization is likely to play an important role in amplifying impressions.

In summary, both proponents of an ecological theory of social perception and evolutionary psychologists propose an adaptive basis of preferences and perceptions, but disagree on the functional significance of specific cues.

2.2. Methodological frameworks

Different methods have been used to determine which facial traits affect perceptions of attractiveness and other social judgments. Some studies were based on ratings of facial morphology. For example, Zebrowitz et al. (1991) used faces rated low and high on babyfaceness to investigate how facial maturity affects hiring recommendations; Rhodes, Chan, Zebrowitz and Simmons (2003) collected ratings of facial masculinity/femininity to test whether facial sexual dimorphism is linked to actual and perceived health.

As ratings of specific aspects of morphology of natural faces might be biased by other facial dimensions, identity, or stereotypical associations (e.g., Komori, Kawamura & Ishihara, 2011; Pound, Penton-Voak & Surridge, 2009), other studies used computer-graphic methods to visually manipulate traits of interest. For example, Rhodes and Tremewan (1996) and Perrett and Penton-Voak (1999) made individual facial images more or less symmetric to test for the effects of symmetry on attractiveness; Perrett et al. (1998) manipulated faces along a male-female shape axis to test for the effects of sexual dimorphism on attractiveness and impressions such as dominance and warmth. While this method made it possible, for example, to visualize the facial correlates of sexual dimorphism, its typical use in two-alternative forced-choice paradigms allows only limited inferences regarding the importance of specific features in judgments of attractiveness and social attributions (see Section 9 of the current thesis).

Yet a different approach took to physically measuring traits. In the context of sexual dimorphism, the size of sexually dimorphic traits such as cheekbone width, eye height or chin length were either used on their own or combined into masculinity indexes to predict attractiveness (Cunningham, 1986; Grammer & Thornhill, 1994; Koehler, Simmons, Rhodes & Peters, 2004; Penton-Voak et al., 2001; Scheib, Gangestad & Thornhill, 1999; Thornhill & Gangestad, 2006; Waynforth, Delwadia & Camm, 2005). One shortcoming of this method is, though, that it requires *a priori* hypotheses about the perceptual determinants of judgments:

if not anticipated, a relationship cannot be detected. A fourth approach circumvents this limitation—data-driven models make it possible to visualize global facial structure that covaries with specific social attributions without the necessity of *a priori* assumptions.

2.2.1. Geometric morphometrics

Quantifying human proportions has a long tradition in physical anthropology and medicine. Physical features have been measured to derive normative data, compare groups (e.g., men and women) and, related to the latter, to identify group-defining characteristics that can be used, for example, to sex skeletal remains. With the advance of statistical and computational methodology, a transition from measuring distances, angles and ratios to geometric morphometric methodology has begun. *Geometric morphometrics* subsume analyses and visualisations that are based on Cartesian coordinates instead of single distances and angles. More specifically, analyses focus on shape: the geometric properties of an object that are invariant to location, scale and orientation (Slice, 2005). Cartesian coordinates of digitised points are transformed into shape coordinates by subjecting them to generalized Procrustes analysis (see, e.g. Slice, 2005, and Section 5), thereby retaining all of the geometric information, i.e. the relative configuration of traits. These methods can be used to derive global face measures of variables such as sexual dimorphism (Komori et al., 2011; Scott, Pound, Stephen, Clark & Penton-Voak, 2010; Valenzano, Mennucci, Tartarelli & Cellerino, 2006), but also for visualising face shape correlates of biological variables such as sex (e.g., Fink, Neave, Manning & Grammer, 2005) and social judgments such as perceived dominance (e.g., Windhager, Schaefer & Fink, 2011) and trustworthiness (Kleisner, Priplatova, Frost & Flegr, 2013). As the number of shape variables often exceeds feasible sample sizes (and for statistical reasons briefly discussed in, e.g., Slice, 2007), principal components analysis (PCA) is often used to reduce the number of variables.

2.2.2. Face space

Parallel to these methods, the psychological literature developed its own computational methods from a somewhat different starting point. The initial impetus came from questions regarding the encoding of faces underlying face recognition. Sirovich and Kirby (1987) laid out how facial images could be described by using principal component analysis. Valentine (1991) then introduced a theoretical model to describe how humans represent faces: "face space" (Turk & Pentland, 1991). O'Toole (2011) summarizes the model as follows: (1) faces are points in a multidimensional space; (2) axes of this multidimensional space describe feature sets that encode individual faces (each face has a specific value on each axis); (3) the Euclidian distance between to faces describes their similarity. Thus, while just one of many tools for morphometricians, PCA also serves a theoretical framework in psychology.

Translating this theoretical model into a "physical face space" (O'Toole, 2011), the first models were purely image- or pixel-based: faces were aligned by the position of eyes, and (grey values of) pixels of each image subjected to a PCA (e.g., O'Toole, Abdi, Deffenbacher & Valentin, 1993). An obvious problem of this approach was that images were not aligned according to feature location, resulting in some face dimensions carrying artefacts due to misalignment of input images. A next step saw shape information removed from faces before subjecting images to a PCA, thus analysing only texture/reflectance information. Craw and Cameron (1991) delineated faces with a set of "control points", averaged the location of these points across faces and then morphed individual images into the average shape. Hancock, Burton and Bruce (1996) further developed this approach by not only conducting a PCA on the "shape-free" (i.e. texture/reflectance) information, but also on the position of the landmark points used to morph individual images into the average shape. They called the resulting PCs "shape vectors".[9] Importantly, this made it possible to separately test for effects of shape and texture on human perception and cognition. Benson and Perrett (1991) and Rowland and Perrett (1995) reconstructed not only the shape of faces, but also their colour information and texture (Tiddeman, Burt & Perrett, 2001), thus making it possible to recreate (and manipulate) more realistic facial images.

Blanz and Vetter (1999) presented a morphable model of 3D face scans, which brought multiple important advancements. First, face shape was independent of lighting in their model. Second, they established correspondence across the whole surface of 3D scans rather than just a landmark template. Third, their their model allowed them to describe individual faces as a trajectory from the average, in terms of both shape and reflectance. Fourth, this morphable model could be fitted to 2D images (Blanz & Vetter, 1999, 2003).

Oosterhof and Todorov (2008) used this morphable model to visualize the facial cues underlying perceptions of trustworthiness and dominance. By regressing judgments of trustworthiness and dominance on principal components of shape, they derived two orthogonal dimensions of face evaluation: valence and dominance[10]. They showed that (a) synthetic faces based on the shape correlates of valence/dominance judgments elicited perceptions of trustworthiness/dominance and (b) that other social judgments such as threat could be reproduced as a function of these two dimensions. Todorov and Oosterhof (2011) also modeled reflectance, and derived six more vectors of social judgments (e.g., attractive, mean, competent), and

[9] While this approach is similar to that used in geometric morphometrics, these shape vectors would not be considered true shape variables in a geometric morphometric framework, because they still contained variance due to differences in scale, as evidenced by face size emerging as the the first principal component in Hancock (2000).

[10] The dominance dimension was rotated to make dimensions of valence and dominance orthogonal.

Todorov, Dotsch, Porter, Oosterhof and Falvello (2013) validated these vectors by showing that they elicit corresponding social judgments.

Similarly, Walker and Vetter (2009) investigated both shape and reflectance correlates of personality, by regressing trait ratings on the faces defining Blanz and Vetter's (1999) 3D morphable model onto PC scores. They then applied the derived vectors to a new set of 2D images: Original images were moved towards shape and reflectance associated with low and high [trait] ratings. These images were then presented in a two-alternative forced choice task, in which participants had to choose the face that looked more like [trait]. For all six traits, participants chose the image manipulated towards higher [trait] above chance, showing that the constructed vectors accurately reflected shape and reflectance associated with the tested trait impressions. Walker, Jiang, Vetter and Sczesny (2011) demonstrated these vectors can also elicit consistent perception of personality cross-culturally.

Said and Todorov (2011) used the Blanz and Vetter face space to build a "statistical model of attractiveness". They generated 4,000 random faces and collected ratings of attractiveness on them. They then regressed ratings of attractiveness on principal components of both shape and reflectance. The resulting model accurately predicted attractiveness of a new set of randomly generated faces (correlation of predicted and rated attractiveness greater than r=.79 for both male and female faces). They also calculated the vector of greatest attractiveness and compared it to the vector of sexual dimorphism (measured as the difference between male and female average). One of the most interesting findings from their model is that male attractiveness is differently related to sexual dimorphism in shape and reflectance: while male masculinity in the reflectance dimensions was relatively similar to the male attractiveness vector, male shape dimensions were closer to the attractiveness vector when more feminine (Said & Todorov, 2011). Similarly, male and female averageness and attractiveness were related differently for shape and reflectance: averageness was found to be attractive especially for shape but not reflectance components.

In summary, recent advances regarding data-driven models provide unprecedented possibilities to visualise and manipulate facial correlates of a range of trait judgments, allowing a more refined investigation of cues used in impression formation.

3. Thesis scope: linking shape to perception

Previous research has made great progress in visually describing face dimensions associated with social perception. While it has been convincingly demonstrated that structural traits resembling emotional expressions are underlying evaluations along the trustworthiness dimension, the dominance dimension has been less explored. That is, it has been found that certain face traits reliably predict perceptions of dominance, and it has been hypothesized that perceptions of dominance are overgeneralized from cues to strength. This was backed by findings that face shape correlates of dominance elicit perceptions of facial maturity and masculinity (Oosterhof & Todorov, 2008). Yet, a strong version of this hypothesis has been under-explored. What are the face shape correlates of masculinity and (actual) strength? Do they indeed link to the perceptions of masculinity and strength (and, ultimately, dominance)?

Thus, one aim of this thesis is to link perceptions back to physical characteristics (sometimes implicitly) assumed to be underlying these perceptions using a data-driven approach. Based on reasoning regarding the adaptive significance of facial cues used in the social evolution of faces, face cues deployed in judgments such as strength should be linked to physical characteristics relevant to these judgments (see Figures 3.0.1 and 3.0.2).

Two of the empirical studies presented in this thesis will link physical characteristics to 3D face shape and test whether face shape correlates of bodily traits inform perceptions of masculinity (Study 1) and strength (Study 3). Another study will investigate the contribution of 3D face shape to facial attractiveness using the example of male facial masculinity (Study 2). Studies on mate preferences have predominantly used two-dimensional stimuli facial photographs. Yet, photographs of the same individual can differ greatly in the information they provide.[1] Even slight head postures change both measurements and perception of faces (e.g., Mignault & Chaudhuri, 2003). Thus, information available from two-dimensional stimuli can be ambiguous, and perhaps even obscuring perceptual effects of subtle facial shape differences. An ambiguity of two-dimensional pictures was, for example, suggested by Swaddle and Reierson (2002), who found that women's preferences for the same male face differed depending on whether the face was presented face-on, or from a lateral view. Swaddle and Reierson concluded that future experiments "need to account for three-dimensional viewing

[1] For unstandardized photographs, variability in attractiveness within subjects can be greater than that between subjects (Jenkins et al., 2011).

of faces if researcher are going to document realistic relations between facial characters and socially (or evolutionarily) important parameters." (Swaddle & Reierson, 2002, p. 2287). The use of three-dimensional stimuli, thus, seems overdue.

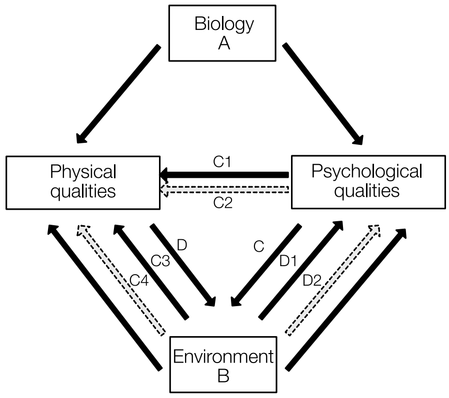
Masculinity and strength were chosen as research topics due to their relevance in both research on mate choice as well as social face evaluation more generally. While Studies 1 and 3 are closer in methodology, the presented experimental studies have been organised according to the aspect they investigated. That is, Part III will present findings on masculinity, while Part IV will focus on strength. Relevant literature will be reviewed at the beginning of the respective sections.

As outlined in the previous section, one of the benefits of using a data-driven approach is that no *a priori* hypothesis regarding the importance of specific traits has to be made. Instead, it can be visualized how face shape is associated with anthropometric traits (e.g., height and weight) and subjective trait-judgments. The latter makes it possible visualize the facial features on which a specific judgment is based on; the former can reveal how bodily characteristics are reflected in the face, and how they might affect social judgments through their reflection in facial appearance.

Many recent studies on social face perception have used FaceGen (Singular Inversions, 2015)[2] to derive shape vectors and create stimuli, which is based on Blanz and Vetter's (1999) morphable model. While FaceGen offers many advantages, especially when it comes to generating tightly controlled stimuli, I used a non-proprietary computer program developed in house, MorphAnalyser (Tiddeman, Duffy & Rabey, 2000)[3]. The use of non-proprietary software allowed me to build (and extend) models from which dimensions can be derived. This also allowed me to integrate additional data on the faces building these models which was necessary to derive face-morphological descriptors of bodily traits such as height and weight.

[2] http://facegen.com

[3] http://cherry.dcs.aber.ac.uk/morphanalyser/version2.4/launch2.4.html



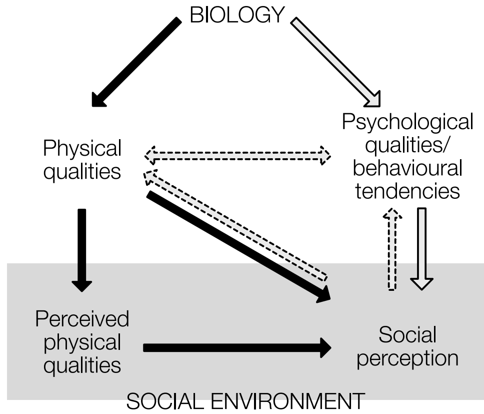


Figure 3.0.1.: Zebrowitz and Collins (1997) proposed four developmental routes that may lead to an association of (facial) appearance and psychological qualities: (A) Common biological causes, i.e. genetic dispositions that affect both appearance and behavioural tendencies. (B) Environmental causes, e.g., a harsh environment leaving physical and emotional traces. (C) Psychological qualities affecting physical attributes. In what is called a "Dorian Gray effect", psychological qualities may lead to a congruent facial appearance (C1, C3). For example, repeated facial expressions may reflect in patterns of wrinkles reinforcing expression-related impressions. Incongruent associations of psychological and physical qualities can arise from efforts to convey desired personality traits by manipulating appearance (C2, C4). (D) Physical attributes affecting psychological qualities, either through a choice of particular social environment, e.g., unattractive people seeking out small gatherings rather than large mixers, or through evoking a particular environment, e.g., unattractive people being viewed and treated as socially awkward. This may lead to self-fulfilling prophecy effects (D1) or self-defeating prophecy effects (e.g., an unattractive person compensating experienced social deficits by becoming more extraverted, D2).

Figure 3.0.2.: The original model by Zebrowitz and Collins was modified by extending it to explicitly include social perception (emphasising that perception and actual attributes are not necessarily congruent), as well as perceived physical qualities which also likely contribute to social perception (e.g., perceived height influences perceived leadership ability, Re et al., 2013).
Black arrows indicate the relationships explored in the current thesis: In order to examine the adaptive significance of facial cues in social perception, biological causes of physical appearance are investigated. This framework can be used to explore how perceptions that have been hypothesized to be rooted in physical qualities can be linked to those physical qualities, although it is not suggested that every perceived trait (and in particular, inferred personality) has a biological cause.
Environmental influences other than social environment (Path B in the original model) are not depicted. Dashed arrows represent paths C and D of the original model and are not further investigated in the current thesis. Note that it is also possible for psychological qualities to affect perceived physical qualities. Sell, Cosmides and Tooby (2014) recently suggested that angry facial expressions serve the function of eliciting perceptions of strength in observers.

Part II.

General Methodology

4. Data sets

The work in this thesis is based on two sets of faces. Face Set 1 is a subset of facial images the Perception Lab has collected over the last 10 years in the course of various research projects. Set 2 was collected as part of the current thesis. Both image sets were collected with the same 3D camera[1] and will be briefly described in the following.

4.1. Face set 1

From a bigger collection of 280 faces, all those images were chosen for which complete scans were available. Incomplete scans can result from problems with viewing angles or poor positioning of the subject, showing in missing areas particularly beneath the chin and around the neck. As the scanner cannot record areas with reflective surfaces, hair is another source of scanning problems, with data often missing from jaws (due to beards, sideburns), or, if no headband is used, ears and forehead. As the sample was unbalanced in terms of ethnicity, I also excluded images of non-caucasian faces.

This left images of 66 women and 66 men. For 59 of these women and 52 men, descriptive data in the form of age, height and weight of the depicted person was available. As all of these variables potentially affect facial perception, I sought to include only those individuals, for which all variables were available. In women and men for which all measures were at hand, body mass index (BMI, $weight[kg]/(height[m])^2$) was heavily skewed (women: skewness=1.71, SE=.31; men: skewness=2.03, SE=.33). Thus, analyses were restricted to a sub-sample; women and men with a BMI above what is considered to be the healthy upper limit (25 WHO, 2013) were excluded[2]. The resulting final sample consisted of 40 women and 40 men and is described in Figure 4.1.1.

4.2. Face set 2

Set 1 was limited in the additional data available for each face. Thus, I collected a second set of data, designed to record more anthropometric variables, as well as other variables that have

[1] `http://www.3dMD.com`

[2] Note that some women had a BMI<18.5 (though not smaller than 18), which is considered slightly underweight.

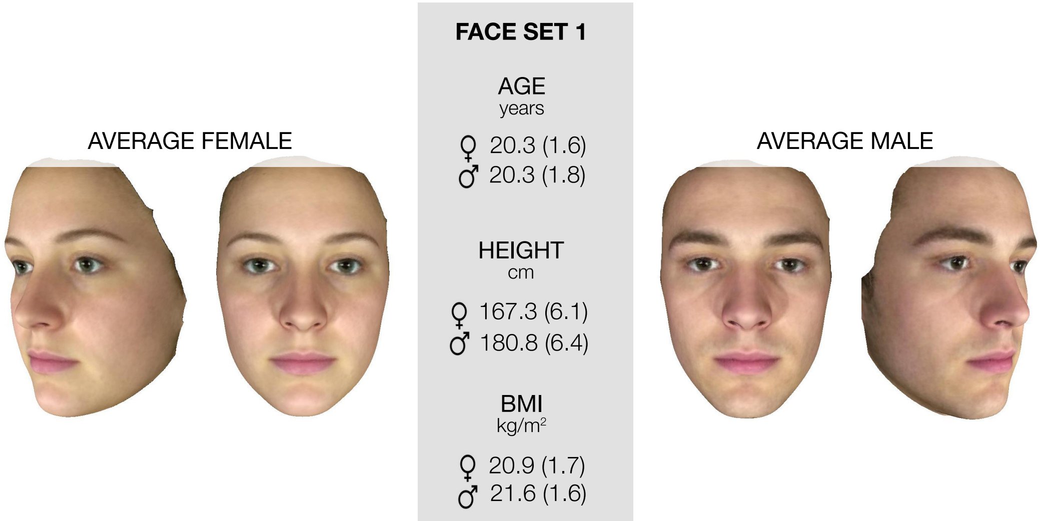


Figure 4.1.1.: Face Set 1 comprised 40 female and male faces each. Values are given as *M(SD)*. See Section 5 on construction of average face images.

been previously linked to masculinity (see below). Data collection took place between January and November 2013. In total, I recruited 140 participants (79 women, M_{age} = 21.1±5.5 years) through the department's study advertising system as well as undergraduate, postgraduate and staff mailing lists.

On first coming in, all participants were asked to use facial wipes. This ensured that faces were bare of make-up for the facial image capturing, and allowed a resting time regarding a potential increase in facial redness caused by the use of the face wipes before taking colour-standardised 2D image. After removing footwear and excess clothing, body height was measured, and weight and body composition (muscle and fat mass) were assessed barefoot using an electrical impedance scale (Tanita SC–330).[3] Left and right hands were scanned for measurements of second-to-fourth digit ratio. Three measures of skeletal frame size were collected: a tape measure was used to record wrist and head circumference, and elbow breadth was taken with a sliding caliper. All three measures were taken three times. Two measures of upper body strength were assessed with a hydraulic dynamometer (Jamar 5030J1, see Part IV for further details on strength measurements). Finally, 3D and 2D images were taken.[4] Participants were asked to remove any jewellery, and their hair was pulled back. Participants were seated at a set distance and relative height to the camera, and asked to maintain a neutral facial expression. A questionnaire on developmental speed, self-reported health, dominance

[3] To account for the weight of remaining clothing, 1 kg was deducted from weight measurements of each participant

[4] To collect colour-calibrated 2D images, the protocol described in Whitehead, Re, Xiao, Ozakinci and Perrett (2012, Experiment 2) was followed.

and self-perceived mating success (see Appendix A) concluded the data collection.

Exclusions were made based on the following criteria: non-caucasian (n=11), poor 3D images (n=5, due to facial hair or technical problems), age (n=6, age<18 or age>mean+3 SD). This left me with a total sample size of 68 women and 50 men (see Figure 4.2.1). No exclusions were made based on BMI.

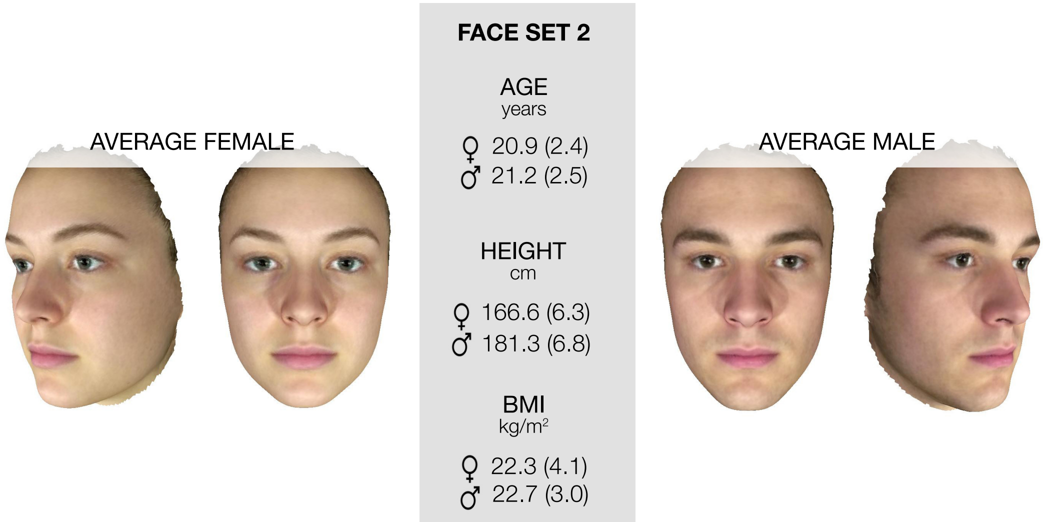


Figure 4.2.1.: Face Set 2 comprised 68 female and 50 male faces. Values are given as *M(SD)*. See Section 5 on construction of average face images.

5. Quantifying face shape

5.1. Preparing faces for analyses

5.1.1. Raw data

The 3D camera provides a surface map of the 3D face structure (a mesh of vertices) and a texture image for each face (see Figure 5.1.1). Thus, shape and texture of face can be manipulated separately.

Figure 5.1.1.: The 3D camera provides separate files for shape (left, surface depth map) and textural information (right).

5.1.2. Delineation

In a first step, facial surface maps were delineated with homologous landmarks. There were several different landmark templates in use when I first joined the lab. The most basic template consisted of 38 landmarks. Most of these landmarks were fairly easy to locate and could thus be delineated with little error, such as the centre of the pupil or tip of the nose. Some landmarks were adopted from 2D delineation templates, such as left and right *zygion*, which denote the greatest width of the face. While these points can be located fairly unambiguously in 2D pictures, in 3D their reliable delineation is more difficult. First, moving the head up and down in a frontal view will change where the face has its greatest width. This is also true for 2D images; yet, as soon as the picture has been taken, in 2D images the location of these points will be fixed. Second, rotating the head from a frontal to a lateral view, it becomes clear that the zygion is actually not a well-defined point but lying on a surface.

In a first step, I thus worked on an operationalised definition of all points that would allow for a reliable positioning of points from different views. Points that could not be operationalised were excluded. For some existing landmarks, position was adjusted; e.g., the landmark that was originally placed at the point of uppermost attachment of the external ear was moved to the point of lowest attachment as the upper part of the ear was often missing for Face Set 1. In addition, landmarks were added on the hairline, jawline and neck to describe the facial outline more comprehensively. This resulted in a total number of 51 landmarks (see Figure 5.1.2 and Appendix B). Faces were delineated in *Morphanalyser 2.3.0* (Set 1) and *Morphanalyser 2.4.0* (Set 2; Tiddeman et al., 2000).

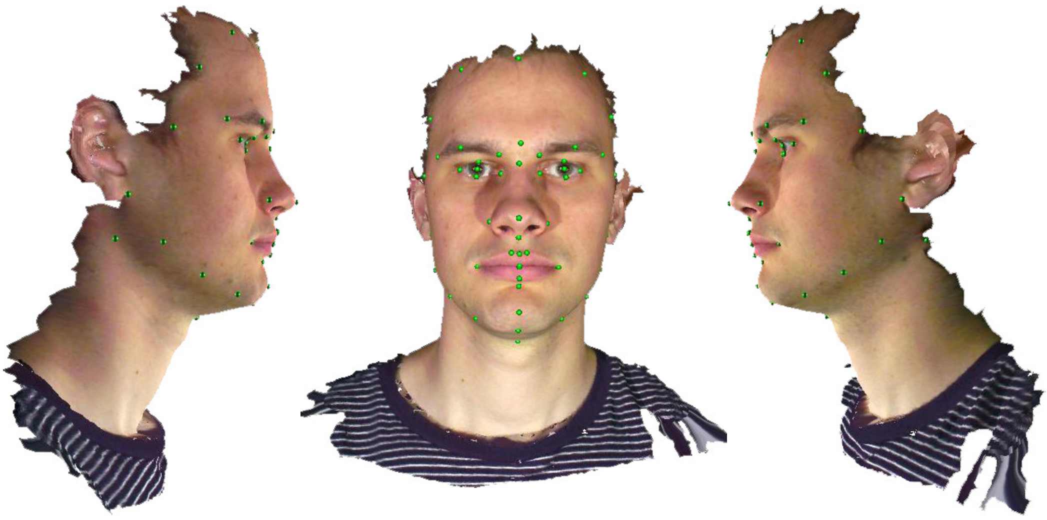


Figure 5.1.2.: Surface maps were delineated with 51 facial landmarks, i.e. homologous points.

5.1.3. Standard head and averaging

To make structures comparable, i.e., accessible to shape analysis, they must be registered in the same shape space. This is achieved by subjecting Cartesian coordinates of delineated landmark points to an algorithm called Generalised Procrustes Analysis (GPA Gower, 1975). General procrustes analyses transforms Cartesian coordinates into shape coordinates: coordinates that reflect variance in shape which is not due to differences in location, orientation, or sized of the original specimens.

Facial images taken with the 3D camera comprise a varying number of vertices. Yet, shape analytic algorithms can only be applied to specimen that have the same number of (corresponding) variables. Thus, an extra step was required to make head models comparable as a whole: Each model must consist of the same number of vertices, and for each model, any particular vertex needs to have a corresponding ("homologous") vertex in each of the other specimens. To achieve this, the landmark templates of all digitised head models were aligned in orientation, rotation and scale using using Procrustes superimposition. In a next step, surface maps of each captured face were re-sampled in accordance to a standard face delineated with the same set of landmark points. Thus, after the alignment process the surface maps of each head model have the same number of tessellations between corresponding landmarks on individual heads (e.g., Re et al., 2011). This establishes homology for the entire facial surface of each head in the set, and allowed me to use whole surfaces rather than landmark templates for further analyses and visualisations. After head models have been brought into the same shape space, *x*-, *y*- and *z*-coordinates of corresponding loci on the facial surface can be averaged across individuals. Maps for the colouration of the facial surface, i.e. image files of the facial texture, can be separately averaged, and brought into alignment with the surface maps (Tiddeman et al., 2000). Figures 4.1.1 and 4.2.1 show average male and female faces of Face Sets 1 and 2.

The standard head that was in use when I first came to the lab was the head of one particular participant. As individual, extreme features could affect the warping process, I created a new standard head in an iterative process. Out of Face Set 1, I chose one female and one male face each, based on the completeness of the scan, a lack of misplaced vertices etc. All female faces were warped to this one female, and a female average face was created; analogously, all male faces were warped to this one male, and a male average face was created. The female average face was warped to the male average face, and then these two faces were averaged to create a new, androgynous standard head, devoid of individual features. Finally, all raw faces were warped to this new standard head.

5.1.4. Masking faces and standardising texture

To eliminate the influence of hairstyle, clothing and cues to strength from neck circumference on perceptual ratings, all 3D heads were masked to show faces only. As colour and textural cues can strongly affect perception (e.g., B. C. Jones et al., 2004; Said & Todorov, 2011; Scott et al., 2010), average male and female face texture images were created using *Psychomorph 4* (Tiddeman et al., 2001). All faces were rendered with this sex-specific standardized texture, so that only face shape differed between same-sex 3D face models. Figure 5.1.3 shows one of the stimulus faces before and after processing.

Figure 5.1.3.: One of the stimuli from Face Set 1 before (left) and after (right) processing, i.e. warping to standard head, masking, and rendering with average texture.

5.1.5. Principal components of shape

Working with the whole surface as opposed to a landmark template considerably increases the amount of dependent variables. Each head model consisted of approximately 11,500 vertices, i.e. was described by about 34,500 *x*-, *y*-, *z*-coordinates. I reduced the dimensions of the data set by subjecting the coordinates defining each head model to a principal component analysis (see Section 2.2), and using the resulting principal component scores as descriptors of face shape.

In both data sets, the first two principal components explained about 30% amount of variance (Face Set 1: 30.7%, Face Set 2: 29.3%). Appendix C visualizes the first six principal components of shape for both sets of faces.

5.2. Describing shape based on group differences

Multiple methods exist in the literature to describe how variables such as attractiveness (e.g., Said & Todorov, 2011) or personality (e.g., Wolffhechel et al., 2014) are reflected in facial shape. For the current thesis, I chose a method that conceptually equates to one of the most frequently used methods in studies of face perception: that is, using the difference in average shape between two groups to describe shape changes between them. For example, to test for the effect of men's masculinity on women's preferences, the difference between men's and women's average face shape has been used to manipulate individual images towards lower or higher masculinity/femininity (Perrett et al., 1998).

While most previous studies have used such vectors to *visually* manipulate individual images, they can also be used to quantify how much an individual face expresses face shape associated with a specific variable. Valenzano et al. (2006) and Komori et al. (2011) recently used this approach to derive an empirical measure of facial sexual dimorphism: based on the average difference between men and women's face shape, they measured sex typicality of individual faces by determining their position on this male to female vector.

In Section 8, I extrapolate these previous methods to calculate a morphological score of sexual dimorphism in 3D faces. Section 8 will also describe how this method can be applied to determine face shape correlates of body height and weight, while Part IV goes on to investigate face shape correlates of body composition, i.e. muscle and fat mass. Appendix D will demonstrate that these measures are relatively stable, and can be cross-validated with an independent set of faces.

Part III.

Male Facial Masculinity

6. Section outline

Men's facial masculinity and its role in mate choice has been a topic of considerable controversy in recent years. This section starts by reviewing literature on women's preferences for men's facial masculinity, and then summarises findings from two empirical studies. Study 1 focusses on explaining perceptions of facial masculinity. Previous literature has shown that objective measures of masculinity only explain a limited amount of variance in perceived masculinity; yet it is pivotal to understand what masculinity entails before making inferences about its significance in mate choice. Study 2 continues the review of the impact of masculinity on women's preference, will test an implicit assumption of many mate choice studies—that masculinity and men's facial attractiveness are related in a curvilinear fashion—and investigates how attractiveness of masculinity changes as a function of different levels of masculinity.

From a methods point of view, Study 1 derives and perceptually validates face-morphological scores of facial sexual dimorphism, body height and BMI, while Study 2 introduces a new testing paradigm for masculinity preferences.

7. Women's preferences for masculinity in men's faces

Trivers (1972) proposed that in species in which females have the higher obligatory investment and males invest in offspring, females should choose males based on their genetic quality, as well as the quality of the parental care males (can) provide. Gangestad and Buss (1993) suggested this may result in a female trade-off between preferences for male investment and male genetic quality. That is, if men differ with regard to heritable fitness, they might demonstrate varying levels of investment in order to be competitive as mates (Gangestad & Buss, 1993). Thus, some men might provide increased levels of investment (but may be providing less heritable fitness to future offspring) while other men might provide higher heritable fitness (but may be investing less in future offspring). Women's preferences for men's facial masculinity have been suggested to reflect this trade-off because masculinity would be indicative of men's heritable fitness. The following section briefly reviews the historical development of this idea.[1]

Thornhill and Gangestad (1993) were the first to suggest that facial masculinity, i.e. exaggerated secondary sexual facial characteristics in men's faces, are attractive to women because they might act as a handicap display: as an honest signal of heritable immunocompetence. Yet, Perrett et al. (1998) found that enhancing masculine facial traits increased both perceived dominance and negative attributions of male faces including decreased quality as a parent (also see Boothroyd, Jones, Burt & Perrett, 2007; Johnston et al., 2001). They concluded that women may differentially prefer characteristics associated with genetic quality and characteristics related to paternal investment. This idea was further developed by Penton-Voak et al. (1999) who found that preferences for masculinity shifted across the female menstrual cycle. In the fertile window of their cycle, women preferred a higher level of masculinity than in the non-fertile phase. As women had also been found more likely to seek extra-pair copulations during the fertile compared to the non-fertile phase of the cycle (Bellis & Baker, 1990), it was suggested that women would pursue a mixed long- and short-term strategy, to

[1] While Gangestad and Buss (1993) as well as the current work focus on women's preferences and female trade-offs in mate choice, it should be noted that men's reproductive success, too, is affected by women's fitness and ability to invest in offspring; men, too, are likely to trade off mate attractiveness and investment (Gangestad & Buss, 1993).

meet a trade-off between less masculine but more agreeable, investing long-term partners and men whose more masculine appearance would indicate good immunocompetence (and thus heritable benefits for potential offspring), but at the same less socially desirable traits (Penton-Voak et al., 1999).

While numerous studies have built on this theoretical framework, the empirical evidence that masculinity acts as an honest signal of immunocompetence has been increasingly called into question (see below). Instead, it has been suggested that masculinity may signal competitive ability to other men, and is only secondarily attractive to women. More recently, it has been suggested that neither of these two approaches fully captures the link between "male testosteronized development and quality" (Gangestad & Eaton, 2013). The different theoretical frameworks will be discussed in the following. Note that for the current thesis it is less important whether masculinity is a signal of a specific aspect of male quality or general condition. Empirical findings, if interpreted somewhat differently by different authors, broadly suggest that women's perception of male attractiveness is sensitive to facial cues of masculinity, and that there might be benefits and costs associated with choosing a (facially) masculine partner (see Section 9).

7.1. Masculinity as a cue to (heritable) health

Numerous studies have linked facial masculinity to perceived (Johnston et al., 2001; Rhodes et al., 2003, 2007; Scott, Swami, Josephson & Penton-Voak, 2008) and (self-reported) actual health (Rhodes et al., 2003; Thornhill & Gangestad, 2006).[2] A link to health has also been suggested based on findings that facial masculinity and symmetry, another potential cue to heritable fitness benefits, as well as preferences for facial masculinity and symmetry are correlated (Little, Jones, DeBruine & Feinberg, 2008; Little, Jones, Waitt et al., 2008).

How could facial masculinity cue health? The immunocompetence handicap hypothesis (ICHH, Folstad & Karter, 1992) proposes that testosterone acts as a handicap (Hamilton & Zuk, 1982; Zahavi, 1975) by suppressing the immune system. Secondary sexual traits would thereby be honestly signaling immunocompetence to potential mates. Male secondary sexual characteristics develop in response to high levels of androgens, but producing and metabolizing testosterone is costly—both energy and metabolites required in the expression of secondary sex characteristics are limited resources (and might lead to higher levels of oxidative stress, e.g., Alonso-Alvarez, Bertrand & Sorci, 2007). However, costs associated with expressing secondary sexual traits are differential—they will be higher for individuals with low (inherited) immunocompetence. "High-quality" individuals, by contrast, can afford

[2] Although Boothroyd et al. (2005) found no evidence that higher facial masculinity was linked to increased perceptions of health.

more costly displays, because disadvantageous effects on health for them will be easier to bear. Thus, immunosuppressant effects of testosterone might reflect an "adaptive" male resource allocation, in that fitness benefits by attracting mates outweigh costs of "drawing energy and essential metabolites" from the immune system (Wedekind & Folstad, 1994).

Although often invoked in explaining preferences for facial sexual dimorphism, evidence for direct effects of testosterone on human immunity is sparse, and a meta-analytic review of cross-species data regards evidence for the ICHH as inconclusive (M. L. Roberts, Buchanan & Evans, 2004, see also Scott et al., 2013). However, testosterone could have indirect effects on the immune system, mediated by glucocorticoids (stress hormones). Chronically elevated levels of glucocorticoids have been found to be immunosuppressive, and correlate in some species with circulating testosterone (M. L. Roberts et al., 2004). This stress-linked version of the ICHH has recently also received some attention in research on humans.

Moore, Cornwell et al. (2011) and Moore, Al Dujaili et al. (2011) found results supporting the stress-linked, but not the original immunocompetence handicap model: women preferred facial cues to low cortisol, and even more so in the fertile phase of their cycle, but they showed no preferences to cues to high levels of testosterone only. In the non-fertile phase, cues to co-occuring high levels of testosterone and low levels of cortisol were found to be more attractive than increased or decreased levels of both hormones (Moore, Cornwell et al., 2011). Cortisol was not found to mediate the relationship between cues to testosterone levels and attractiveness; rather, testosterone was moderating the effect of cortisol on attractiveness: under high co-occuring levels of testosterone, cues to cortisol had a smaller effect on attractiveness ratings (Moore, Al Dujaili et al., 2011). Moore and colleagues suggested that instead of testosterone, it might be cortisol that acts as cue to a heritable component of health, whereas "testosterone might be a proxy of male ability to cope efficiently with stressors" (Moore, Al Dujaili et al., 2011, p. 269). Rantala et al. (2012) reported further evidence for the stress-linked ICHH: higher levels of testosterone were linked to both a higher immune response to a hepatitis B vaccine and higher facial attractiveness, and these relationships were stronger in men with lower levels of cortisol.[3]

Boonekamp, Ros and Verhulst (2008), too, suggested a trade-off between immunocompetence and sexual signaling, but in an opposite way to that proposed by the ICHH: immune activation would lead to suppressed testosterone levels. Repeated, or endured response to immune threats might thus lead to less masculine phenotypes (Puts, Jones & DeBruine, 2012). Although this theory is more compatible with findings by M. L. Roberts et al. (2004), it is less

[3]Note that a later study by Rantala et al. (2013) suggested the relationship of immunocompetence and facial attractiveness was likely to be mediated by facial adiposity rather than masculinity. However, masculinity was not objectively measured but based on perceptual ratings—as Study 1 will show, perceived masculinity is a suboptimal measure of sexual dimorphism in facial shape.

compatible with the findings by Rantala et al. (2012).

7.2. Masculinity as a cue to status/dominance

Recently, it has been suggested that instead of having primarily evolved under inter-sexual selection pressures (i.e. for cueing health to women), men's facial masculinity may have primarily evolved under intra-sexual selection pressures (Scott, Clark, Boothroyd & Penton-Voak, 2013). Men's masculinized voices are perceived to belong to more dominant men (e.g., Puts et al., 2012), and enhancing facial masculinity of male (and female) photographs increases perceived dominance (Boothroyd et al., 2007; DeBruine et al., 2006; Perrett et al., 1998; Swaddle & Reierson, 2002; Watkins, Jones & DeBruine, 2010; Watkins, Quist, Smith, DeBruine & Jones, 2012; Windhager et al., 2011). Male facial robustness might have also direct benefits in male-male competition—the greater robustness of male skulls may be related to higher levels of physical violence in men (Puts, 2010; Puts et al., 2012); Stirrat, Stulp and Pollet (2012) showed that wider-faced men were less likely to die from male-male physical violence.

Does facial masculinity, reflecting high levels of testosterone, reliably signal dominance, and thereby status? According to Mueller and Mazur (1997), it does. In a sample of cadets, facial dominance was related to higher status (i.e., military rank) attained. Swaddle and Reierson (2002), too, found that faces (manipulated to) showing putative markers of increased testosterone were perceived as more dominant. This association of masculinity and perceived and actual dominance might be mediated via two routes: physical strength, and a drive to seek for status.

Physical Strength Facial masculinity correlates positively with measures of hand grip strength (Fink, Neave & Seydel, 2007; Windhager et al., 2011), a proxy of general body strength (briefly reviewed in Wind, Takken, Helders & Engelbert, 2010). In addition, facial shape has been found to correlate with second-to-fourth (2D:4D) digit ratio (Fink et al., 2005), which is thought to reflect early organising effects of testosterone on both physical abilities and facial shape—the higher the prenatal testosterone exposure, the lower the 2D:4D ratio. Second-to-fourth digit ratio correlates positively with both male strength and physical performance in men and women (e.g., Fink, Thanzami, Seydel & Manning, 2006). Moreover, a lower 2D:4D digit ratio is associated with more masculine and a higher 2D:4D with more feminine facial features (Fink et al., 2005).

Status-seeking Whereas early studies suggested a causal relationship between testosterone and aggressive behaviour, many studies indicate that testosterone might be linked to status-seeking behaviour (which at times might require higher levels of aggression; e.g., Rowe, Maughan, Worthman, Costello & Angold, 2004; Schaal, Tremblay, Soussignan & Susman, 1996). Increased testosterone levels in response to anticipation and success in competitive tasks have been interpreted as evidence that testosterone might mediate dominant behaviour intended to gain and maintain status (e.g., Mazur & Booth, 1998). After winning a competition, men with more masculine faces show higher increases in levels of circulating testosterone than men with less masculine faces (Pound et al., 2009), and the willingness to engage in another competitive task after defeat is greater for men with post-experimentally higher levels of testosterone (Mehta & Josephs, 2006). Similarly, a supposedly testosterone-dependent trait (facial width-to-height ratio), that was previously thought to be linked to aggression, has been recently found to predict achievement drive (Lewis, Lefevre & Bates, 2012).

7.3. Masculinity in the context of life history theory

While recent discussions have focussed on an either-/or-approach (masculinity is either a cue to health or dominance), there is a broader theoretical context that may offer an integrative framework for previous findings (Gangestad & Eaton, 2013): testosterone might mediate the allocation of male resources to survival vs reproductive effort, as well as parenting vs mating efforts (e.g., Ellison, 2003; Muehlenbein & Bribiescas, 2005).

To maximise their inclusive fitness, organisms need to optimally allocate their finite resources to competing functions, such as reproduction, maintenance, storage and growth, and thus face a range of unavoidable trade-offs (Muehlenbein & Bribiescas, 2005). Among the most important of all trade-offs is that of reproductive effort and survival: investing in the former limits the available resources to the latter, and thus implies costs to survival or future fertility (Ellison, 2003; Muehlenbein & Bribiescas, 2005).

While men's reproductive effort is small in terms of obligatory minimal investment in offspring (male gamete production requires little energetic investment), energy devoted to reproductive effort in the form of mating effort such as mate guarding, sexual activity and intrasexual competition can be considerable. Male competitive ability is facilitated by physical strength/muscularity. According to Ellison (2003), the metabolically expensive production and maintenance of muscle form can be considered somatic reproductive effort. Evidence suggests that this somatic reproductive effort is modulated by testosterone: energy constraints can lead to the suppression of testosterone levels in order to allocate resources to other body functions such as growth or immune function. Testosterone not only contributes to the regulation of somatic reproductive effort but may also help to regulate behavioural mating effort, by impacting on libido and possibly also on confidence and assertiveness in competitive social interactions. Indeed, high levels of testosterone have been shown to facilitate mating effort: men with higher testosterone levels have been found to seek out more mating opportunities (e.g., Pollet, der Meij, Cobey & Buunk, 2011) and show lower levels of commitment (e.g., Booth & Dabbs, 1993). In contrast, low levels of testosterone may facilitate parenting effort: lower levels of testosterone have been associated with increased spousal investment and higher commitment to romantic relationships (Burnham et al., 2003; Gray, Kahlenberg, Barrett, Lipson & Ellison, 2002). Facial masculinity may thus be a visual indicator of men's general condition, as well as male pursuit of mating vs parenting effort strategies.

In summary, testosterone might affect female mate choice as (1) testosterone could be related to heritable immunity, but evidence for a link of heritable immunity and testosterone is weak; (2) testosterone could be related to intra-sexual competition; (3) testosterone could be more generally related to mating effort and somatic signs of reproductive effort (such as bulk and muscle mass, see Section 8 and Part IV of the current thesis). Yet, women face a trade-off, as high levels of testosterone are associated with reduced investment in offspring, higher levels of infidelity and other non-desirable social traits. This trade-off is the subject of Section 9.

8. Study 1: What makes men's faces masculine?[1]

8.1. Introduction

A plethora of published studies has examined the role of men's facial masculinity on mate choice and interpersonal judgments such as leadership or dominance (e.g., Little, Apicella & Marlowe, 2007; Perrett et al., 1998). Despite this, studies to date have failed to provide a clear and comprehensive account of what constitutes a "masculine" face when non-shape cues such as skin texture and facial hair are controlled for. Measures of sexual dimorphism in face shape have been found to account for as little as 6–11% of variance in ratings of masculinity (Koehler et al., 2004; Komori et al., 2011; Pound et al., 2009). Moreover, although ratings of masculinity have been linked to judgments of attractiveness (Koehler et al., 2004; Rhodes et al., 2005, 2007; Scott et al., 2010), several studies have failed to find a relationship between morphological masculinity and attractiveness (Koehler et al., 2004; Penton-Voak et al., 2001; Scott et al., 2010; Stephen et al., 2012; Thornhill & Gangestad, 2006; Waynforth et al., 2005). This has led some researchers to conclude that morphological measures of masculinity are not overly useful, as they fail to capture masculinity as perceived by raters. Others have reasoned that perceptual ratings of masculinity are problematic, as they appear to be confounded by unknown parameters. Komori et al. (2011) termed these parameters "sex-irrelevant characteristics" and suggested they reflect sexual stereotypes of personality.

Here, we tested a different hypothesis. Given the sexual dimorphism in body height and weight of men and women (e.g., Gaulin & Boster, 1985), we investigated whether facial correlates of these variables affect the perception of men's masculinity. Height and weight affect face structure (e.g., Coetzee, Chen, Perrett & Stephen, 2010; Mitteroecker, Gunz, Windhager & Schaefer, 2013), and the resultant facial cues may affect not only the perception of body size but also masculinity.

Some researchers have challenged the validity of using 2D photographs in studies assessing the perception of gender, since 2D images do not fully depict the prominence of features that

[1] Part of the work in this section has been published in *Perception*. The reference for this work is as follows: Holzleitner, I. J., Hunter, D. W., Tiddeman, B. P., Seck, A., Re, D. E., & Perrett, D. I. (2014). Men's facial masculinity: when (body) size matters. *Perception, 43*, 1191-1202.

differ between men and women (e.g., eyebrow ridge or jaw protuberance; Bruce et al., 1993; Burton, Bruce & Dench, 1993, see also Swaddle & Reierson, 2002). Valenzano et al. (2006) and Komori et al. (2011) described an objective score of sexual dimorphism based on the average difference between men and women's 2D face shape. We extrapolated these previous methods to calculate the morphological masculinity of 3D faces. In addition, we calculated the average shape difference between short and tall men, as well as the average difference between men with a low and high body mass index (BMI), to test whether facial correlates of body height and weight predict perceptions of height and weight. Furthermore, we examined whether facial cues to height and weight may account for previously unexplained variance in ratings of masculinity. We hypothesized that within a healthy weight range both morphological cues to weight and height would be positively associated with the perception of facial masculinity.

Study 1 comprised two parts: In Study 1.1, we calculated the measure of morphological masculinity, linked this measure to perceived masculinity and visually compared similarities and differences between masculine facial architecture and perception of masculinity. In Study 1.2, we derived the facial correlates of height and BMI, and then tested the hypothesis that facial cues to these parameters contribute to the perception of masculinity.

8.2. Study 1.1: Perceived and morphological masculinity differ

8.2.1. Methods

8.2.1.1. Stimulus data set

The stimulus set comprised the faces of Face Set 1 (see Section 4.1).

8.2.1.2. Morphological masculinity

After head models were subjected to a principal component analysis, the resulting 80 principal component (PC) scores served as the computational basis for morphological masculinity (see Section 5). Average male scores and the (androgynous) sample average of each component were computed. We defined the morphological masculinity axis as the direction from the androgynous sample average $\vec{a}$ to the male average $\vec{m}$ (arrows are denoting vectors). We calculated individual masculinity scores (MS) as the distance along the morphological masculinity axis from the component scores for a subject's face $\vec{i}$ to the point on the morphological masculinity axis closest to the male average component scores $\vec{a}$. This distance was then normalized by dividing by the magnitude of the masculinity axis (i.e. the distance between male and sample average) to ensure that androgynous faces receive a score of 0, and faces with average masculinity receive a score of 1. That is, each individual face $\vec{i}$ was projected onto the morphological masculinity axis $\vec{m} - \vec{a}$ using:

$$MS(\vec{i}) = \frac{(\vec{i} - \vec{a}) \cdot (\vec{m} - \vec{a})}{\|\vec{m} - \vec{a}\|^2}$$

where $\|\,\|$ gives the magnitude (length) of the vector. Figure 8.2.1 visualizes facial correlates of this shape vector.

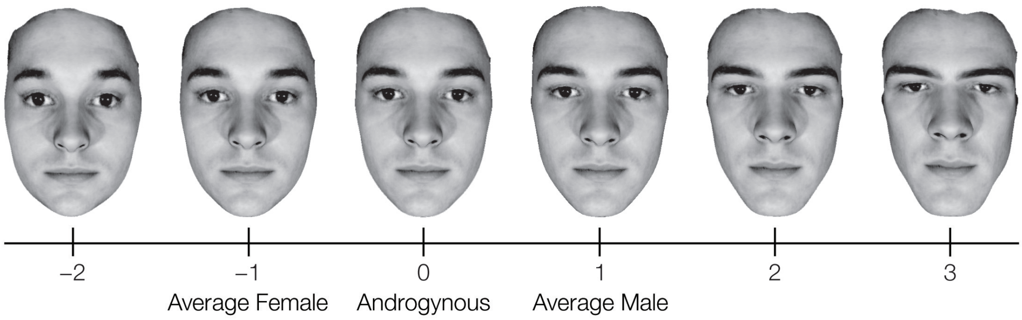


Figure 8.2.1.: Masculinity scores were based on the shape difference between the (androgynous) sample average (0) and the average male face (1). The hypermasculine faces with scores 2 and 3 illustrate changes in facial shape along this vector towards higher masculinity and were generated by applying 200% and 300% of the difference between the androgynous and the average male face to the androgynous face, respectively. The hyperfeminine face with score –2 visualizes changes towards higher femininity (all faces were rendered with the same skin texture for illustration purposes). Individual masculinity scores in the 40 men of Face Set 1 ranged from –0.4 to 2.1.

To test whether the calculated scores were indeed detecting morphological differences related to sex, we also calculated the morphological masculinity of female faces, and both female and male scores were employed in a discriminant analysis. The resulting discriminant function yielded correct sex classifications for 92.5% of the faces (Wilks' λ=.264; df=1; χ^2=103.3, p<.001).

8.2.1.3. Perceived masculinity

Twenty Caucasian female students (M_{age} = 21.5±2.5 years) from the University of St Andrews rated the masculinity of the faces on a 7-point Likert-type scale (1-*not masculine at all*, 7-*very masculine*). Prior to the rating, participants were presented with frontal 2D images of all face models to provide an overview of stimulus variability. The 3D face stimuli were presented on a computer screen in randomized order. They were rotated from –50° to +50° from left to right while simultaneously being rotated from –15° to +15° up and down, resulting in the stimuli "bobbing" in a sinusoidal manner. Images were presented individually against a black background and remained visible until a rating was made[2]. All procedures were approved by the University of St Andrews Teaching and Research Ethics Committee.

Inter-rater reliability for ratings of masculinity was high (Cronbach's α=.82). Masculinity ratings by the 20 participants were averaged for each of the 40 faces. Regression analyses were used to test the predictive value of morphological scores for perceptual judgments

8.2.1.4. Similarities and differences of morphological and perceived masculinity

To visualize the similarities and differences in facial structure between perceived and morphological masculinity, composites of the 10 faces scoring lowest and highest on each variable were generated (see Table 8.1). The difference between low and high perceived masculinity composite faces was calculated and translated into standard deviation (SD) units for perceived masculinity observed in the sample (the difference between low and high prototypes in ratings on the 7-point scale was 1.3, equating to 2.5 SD). To visualize the face shape associated with a perceived masculinity of 5 SD below the mean and 5 SD above the mean, two times the difference between low and high prototypes was subtracted from or added to the average male face. Analogously, the transform amount equivalent of 5 SD of morphological masculinity was subtracted and added from the average male face to create transforms reflecting the face shape associated with low and high morphological masculinity.

Table 8.1.: Perceived and morphological masculinity in Face Set 1. The difference in shape between composite head models of the 10 men scoring lowest and highest on perceived and morphological masculinity served to visualize face shape associated with masculinity. Values are given as *M(SD)*. Perceived masculinity was rated on a 7-point Likert-type scale; morphological masculinity describes how close a face was to the androgynous sample average (0) as compared to the average male face (1).

	Prototype	N	Perceived masculinity	Morphological masculinity
Perceived masculinity	Mean	40	4.34 (0.52)	1.01 (0.63)
	Low	10	3.67 (0.29)	0.79 (0.63)
	High	10	4.98 (0.21)	1.21 (0.42)
Morphological masculinity	Mean	40	4.34 (0.52)	1.01 (0.63)
	Low	10	4.11 (0.40)	0.21 (0.38)
	High	10	4.49 (0.51)	1.79 (0.21)

[2]A demo of the rating interface used in the empirical studies in this thesis can be found at `http://perceptionlab.com/~Iris/SI.html`.

8.2.2. Results

Morphological masculinity significantly predicted perceived masculinity (β=.33, R^2=.11, F(1,38)=4.57, p=.039, see Figure 8.2.2). Figure 8.2.3 visualizes male face shape associated with perceived and morphological masculinity.

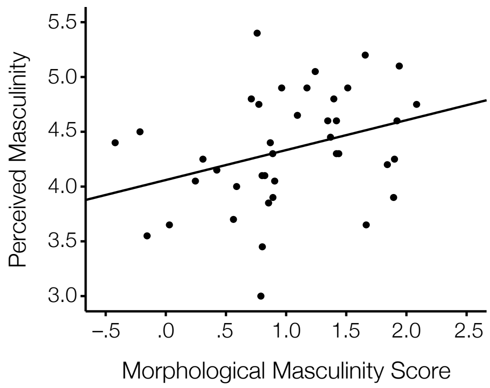


Figure 8.2.2.: Morphological masculinity was calculated based on shape differences between the (androgynous) sample average and the average male face (denoted by 0 and 1 on the x-axis), and was a moderate predictor of perceived masculinity (R^2=.11, N=40).

As can be seen from Figure 8.2.3, facial correlates of morphological and perceived masculinity showed similarities as well as differences.

Compared to men rated to look less masculine, highly masculine men were found to have more robust-looking faces: the forehead was less bulbous, the eyes smaller, the nose bridge and mouth wider, and eye brow ridges, chin and jaw were more pronounced. Overall, highly masculine men appeared to look heavier and more mature.

Men having a face shape more masculine than the male average, compared to men with a less than average male face shape, too, appeared to have a more robust faces with a less bulbous forehead. Again, brow ridges, jaw and chin increased in prominence with increasing masculinity. However, in contrast to perceived masculinity, morphological masculinity did not seem to be associated with changes in eye size, and changes in the chin and jaw region were less pronounced compared to those associated with perceived masculinity. In addition, morphologically masculine men did not necessarily appear heavier, but had wider faces with more pronounced zygomatic arches (cheekbones).

8.2.3. Discussion

Study 1.1 extended a previously used objective measure of facial masculinity to 3D shape and related it to perceived masculinity. In line with our prediction, morphological masculinity

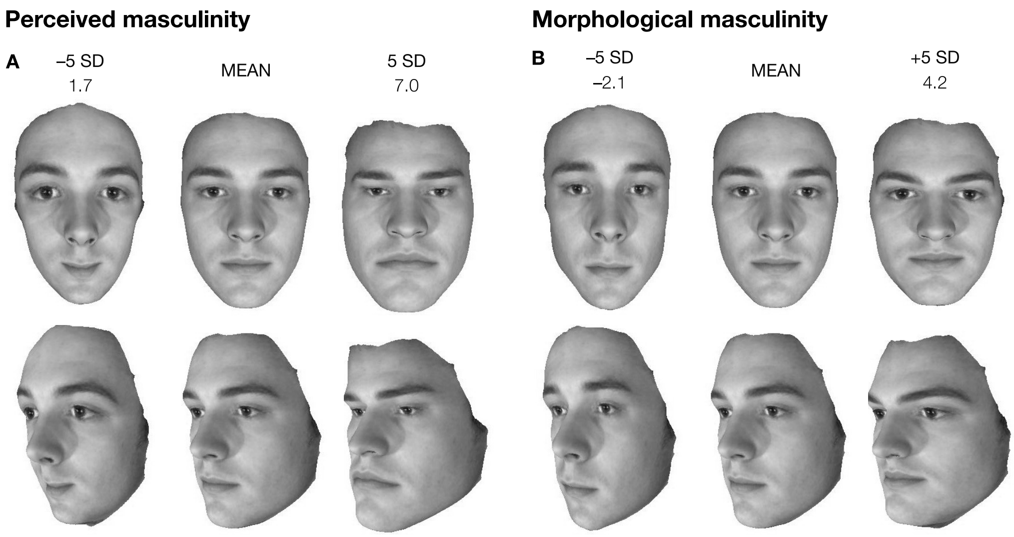


Figure 8.2.3.: Face shape associated with perceived (A) and morphological (B) masculinity. Visualizations reflect face shape associated with ±5 SD of perceived and morphological masculinity, based on the difference in face shape between the 10 men scoring lowest and highest on perceived masculinity, and lowest and highest on morphological masculinity (see Table 8.1). Please note that the transform amount of ±5 SD was chosen to increase the salience of changes and goes beyond what would be observed in natural faces.

predicted perceived masculinity, albeit to a limited extent. In a next step, we investigated which facial traits drive the perception of masculinity, and how these traits may differ from differences between the average male and female face shape, i.e. sexual dimorphism.

Both morphological and perceived masculinity were found to be associated with facial robustness and changes in the ratio of upper- and mid-face height to lower face height. Robust, short and wide faces were perceived as more masculine; elongated and more gracile faces as less masculine. This concurred with shape differences based on low and high scores of morphological masculinity, and fits with traits previously found to be sexually dimorphic. For example, men have larger airways than women, reflecting a higher energy expenditure due overall greater body size and muscle mass (Enlow & Hans, 1996). As airways are a developmental key stone in facial development, this leads to a principal difference in men and women's faces: On average, male noses are larger, wider, with a straight to convex shape and the nose tip pointing downwards. Male noses are more protrusive, which also models the forehead contiguous with the nose into a more protrusive position, and leads to a typically more sloping forehead in men, with protrusive supraorbital and globular part (Enlow & Hans, 1996). Women, on the other hand, have a smaller frontal sinus, which makes the temporal regions of their lateral forehead appear less full and more sloping. Overall, the female face appears flatter, proportionately wider and more gracile—which in the current study was reflected in face shape correlates of both perceived masculinity as well as our measure of sexually dimorphic face shape.

However, as reflected by the moderate association of morphological and perceived masculinity, not all perceptually dimorphic traits corresponded to sexual dimorphism as captured by the average difference in male and female face shape. First, perceived masculinity was linked to eye size, with big eyes being perceived as less masculine. Eye size has previously been found to be greater in women than men (Mitteroecker, Windhager, Müller & Schaefer, 2015; Penton-Voak et al., 2001; Thornhill & Gangestad, 2006; Waynforth et al., 2005, but see Koehler et al., 2004). Yet, prototypes based on scores of morphological masculinity showed no changes with eye size. As I did not test for statistical differences *between* the sexes, it could be that eye size was sexually dimorphic in the tested sample although this difference could only be small (or else would show in the visualisations of face shape associated with morphological sexual dimorphism). Nonetheless, perceptually, eye size seems to be an important feature when judging masculinity. Two explanations seems possible: either raters overgeneralize sexually dimorphic features when making judgments of masculinity; or it could be that eye size is "mistaken" as a sexually dimorphic feature when it is actually related to height/overall size. Eyeball size scales with negative allometry relative to body size, while orbits scale with strong positive allometry to facial size (Lieberman, 2011)—that is, being shorter may lead to larger(-looking) eyes, and as women are on average shorter than men, larger eyes might be associated with being female. Indeed, the composite faces of perceived masculinity showed a trend to differ in height: men in the low perceived masculinity composite face showed a trend to be shorter (179.6 cm) than men in the high perceived masculinity composite face (183.4 cm; t=1.54, p=.140), while men in the low and high morphological masculinity composites were closer in height (t=.59, p=.562; the mean height of the low composite face was actually marginally greater than that of the high morphological masculinity composite).

Second, perceptual but not morphological masculinity was associated with changes in facial expression. Perceptually highly masculine faces had a much more "unfriendly" expression than those faces perceived the least masculine, which showed a subtle smile. While facial expression themselves are not sexually dimorphic (both men and women smile), smiling might be perceived as a more feminine trait (Hess, Adams Jr., Grammer & Kleck, 2009; Hess, Adams Jr. & Kleck, 2009), and thus interact with the association of actual and perceived sexual dimorphism.

Third, both morphological as well as perceptual masculinity were associated with a more robust facial appearance, but differences between low and high levels of perceptually masculinity were more pronounced than those between low and high levels of morphological masculinity. Faces rated the least masculine appeared notably slimmer and more gracile than those rated as highly masculine. Men are on average taller and heavier than women, and both body weight and height have been previously shown to affect facial shape (e.g., Coetzee et al., 2010; Mitteroecker et al., 2013). It seems likely that facial cues to these characteristics may contribute to the perception of sexual dimorphism/masculinity and this formed the rationale for Study 1.2.

8.3. Study 1.2: Facial cues to body size contribute to the perception of masculinity

In line with previous findings, Study 1.1 showed a discrepancy between masculinity of face shape and perception thereof. Study 1.2 investigated whether facial cues to body height and weight may explain some variance in perception of masculinity previously unaccounted for by morphological masculinity.

8.3.1. Methods

8.3.1.1. Face-morphological height and BMI scores

Analogously to the calculation of morphological masculinity (see Section 8.2.1.2), morphological scores were calculated separately for height and BMI. Average PC scores were calculated for short and tall men, as well as men with high and low BMI (see Table 8.2), and the resulting shape vectors were used to assign each face a score on facial correlates of height and BMI, respectively (see Figure 8.3.1 for visual representations of the two vectors). Men in the low and high BMI groups were matched so they did not differ in height (mean difference=1.8 cm, $t(17)=0.67$, $p=.520$); likewise men in the low and high height groups were matched so they did not differ in BMI (mean difference=0.04 kg/m^2, $t(15)=0.05$, $p=.959$). Resulting morphological scores of height and BMI were not correlated (Pearson's $r(40)=-.16$, $p=.316$). Morphological height and BMI scores were cross-validated with the independent Face Set 2 in Appendix D.

8.3.1.2. Experimental validation of face-morphological scores

Stimuli Stimuli were the 40 male 3D face models from Face Set 1.

Procedure *Perceived Weight* Seventeen students (2 male, age 21.6±3.6 years) from the University of St Andrews rated the body weight of persons depicted in the face stimuli on a 7-point Likert-type scale from 1-*Very underweight* to 7-*Very overweight*. The experimental set-up was the same as in Study 1.1 (see Section 8.2.1.3).

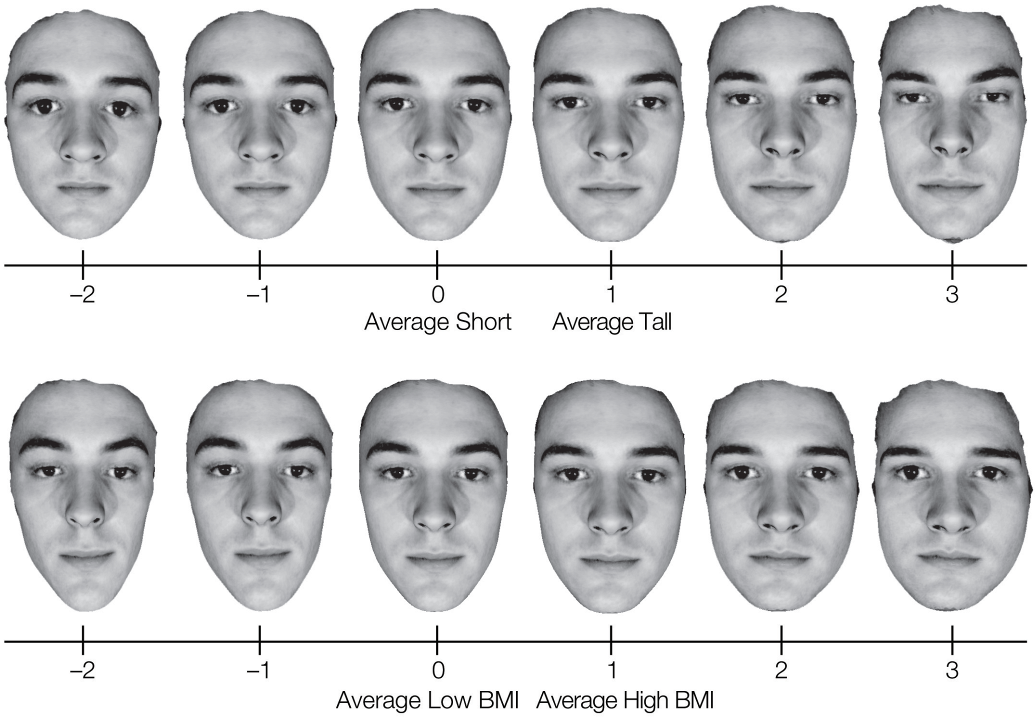


Figure 8.3.1.: Morphological height and BMI scores were based on the shape difference associated with short (0) and tall (1) height (top row), and low (0) and high (1) BMI (bottom row). The synthetic faces –2 and –1 illustrate shape changes towards lower height/BMI, while the synthetic faces 2 and 3 illustrate changes towards higher height/BMI. Individual height scores in the sample ranged from –1.3 to 2.1, while BMI scores ranged from –0.9 to 3.1.

Table 8.2.: Computing facial correlates of BMI and height. Face-morphological scores of height and BMI were based on the differences in face shape between short and tall men, and men with low and high BMI, respectively. Values are given as *M(SD)*; signficant differences (p<.05) between low and high prototypes are indicated in bold. Note that due to constraints in sample size in constructing BMI-controlled height scores and height-controlled BMI scores the number of faces that defined the height and weight vectors differ.

	Prototype	N	Height (cm)	BMI (kg m^{-2})
Height	Mean	40	180.8 (6.4)	21.6 (1.6)
	Low	9	**171.9 (3.2)**	21.6 (1.2)
	High	8	**189.0 (2.1)**	21.5 (1.9)
BMI	Mean	40	180.8 (6.4)	21.6 (1.6)
	Low	9	180.2 (5.8)	**19.3 (0.7)**
	High	10	182.0 (6.0)	**23.5 (0.4)**

Perceived Height Thirty-nine participants (12 male, age 26.8±10.1 years) rated body height of depicted persons on a 7-point Likert-type scale from 1-*Very short*, to 7-*Very tall* in an online study.

For all studies, informed consent of participants was obtained either in written form or electronically. All procedures were approved by the University of St Andrews Teaching and Research Ethics Committee. Inter-rater reliability was high for both ratings (Cronbach's α perceived weight=.92, perceived height=.93).

8.3.2. Results

Ratings of masculinity, height and weight were averaged separately for each of the 40 faces. Regression analysis was used to test the predictive value of morphological scores for perceptual judgments.

The morphological height score was a strong predictor of perceived height (β=.42, $R^2$2=.18, F(1,38)=8.04, p=.007), whereas actual height was not related to perceived height ($R^2$2=.01, F(1,38)=0.52, p=.475; Figure 8.3.2, top row). The morphological BMI score was a strong predictor of perceived weight (β=.65, R^2=.43, F(1,38)=28.13, p<.001). Actual BMI was not a significant predictor of perceived weight, although it showed a trend for a relationship in the expected direction (β=.26, R^2=.07, F(1,38)=2.67, p=.110; Figure 8.3.2, bottom row).

Adding the morphological height and BMI scores to the regression model predicting

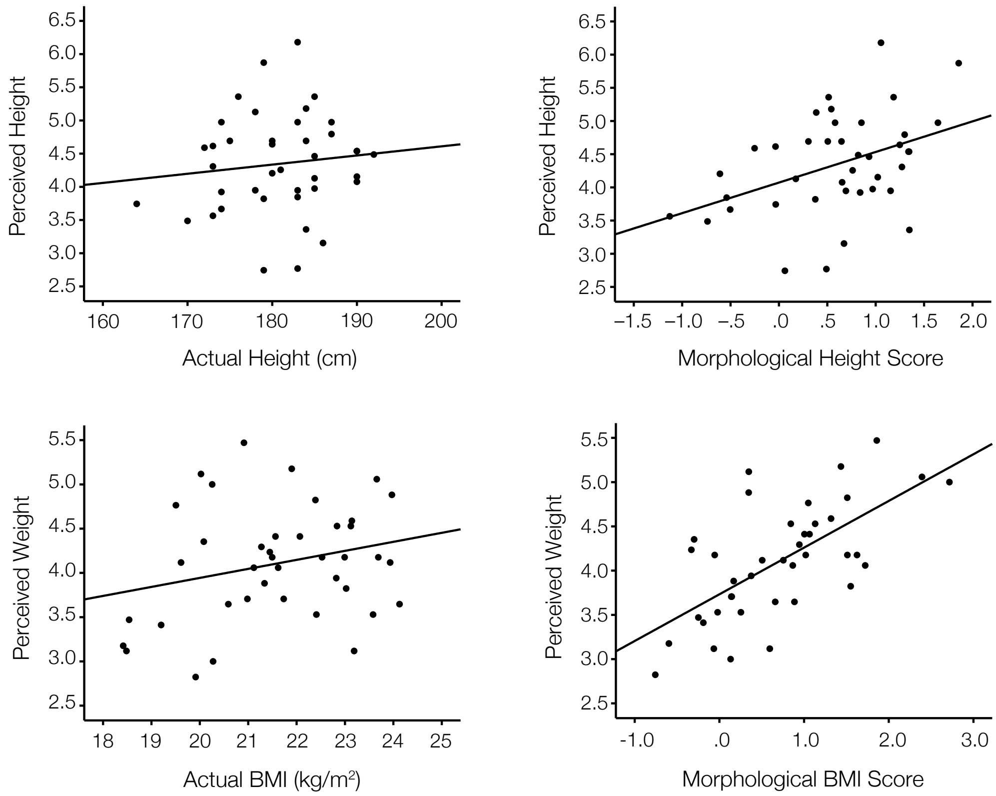


Figure 8.3.2.: Correlations of actual height/BMI, ratings of height/weight, and corresponding face-morphological scores and ratings (N=40). Face-morphological scores were better predictors of perceived height and weight than actual height and BMI (height: actual R^2=.01 vs morphological scores R^2=.18; BMI: actual R^2=.07 vs morphological scores R^2=.43).

perceived masculinity increased the variance explained from 11% for masculinity scores as sole predictor (AIC=−54.9) to 34% (R^2=.34, F(3,36)=6.18, p=.002; AIC=−61.5). All three morphological scores were found to be significant predictors of perceived masculinity, with morphological masculinity scores being the strongest predictor (β=.46, p=.006), followed by morphological scores of height (β=.42, p=.011) and BMI (β=.35, p=.016; tolerance for all variables>.75; variance inflation factor<1.3). While morphological masculinity was not significantly related to facial cues to BMI (Pearson's r(40)=.20, p=.210), it was significantly negatively correlated with facial cues to height (Pearson's r(40)=−.49, p=.001).

Brand and Bradley (2012) have argued that the use of average-based rating scores inflates effect sizes (but see McCormick, 2013). We also examined the association of morphological variables and masculinity ratings of individual raters. We compared the average fit of models using only morphological masculinity as a predictor of individually perceived masculinity (simple models) with those using all three morphological scores (masculinity, height and BMI) as predictors (full models). Full models (mean R^2=0.14, mean adjusted R^2=0.07, mean AIC=126.4) were not found to explain substantially more variance than simple models (mean R^2=.04, mean adjusted R^2=0.01, mean AIC=127.4) as indicated by the minor change in the mean AIC. Thus, observed effect sizes in the group-based analysis were indeed bigger than in the individual-based analysis, but pointed in the same direction (McCormick, 2013).

8.4. Discussion

Our results show that it is possible to derive meaningful morphological scores of body height and BMI from 3D face shape. Face-morphological scores of height and BMI strongly predicted perceived height and weight. Though morphological masculinity alone moderately predicted perceived masculinity of men's colour- and texture-standardized faces, morphological correlates of height and BMI made additional and independent contributions to the perception of men's masculinity.

The physical characteristics that influence the perception of masculinity have proven remarkably elusive. Like others, we find that sex differences in face structure explain only 11% of perceived masculinity (6 - 11% in Koehler et al., 2004; Komori et al., 2011; Pound et al., 2009). Since morphological masculinity predicted gender in our sample correctly for 92.5% of faces, we suggest that the weak relationship between morphological and perceived masculinity cannot be explained by an inadequate structural estimation of sexual dimorphism. In line with others, we propose it is the perception of masculinity that is poorly understood. Whereas Komori et al. (2011) explain some of the discrepancy of morphological and perceptual masculinity with social stereotypes of personality, the aim of the current study was to investigate why specific social perceptions may be driven by certain face shape features. In particular, we tested whether face shape correlates of body dimensions impact on the perception of masculinity. We used a simple computational method to show that face structure associated with quantitative anthropometric variables such as body height and BMI affects the perception of facial masculinity. Men are perceived as masculine not only based on how much their face shape differs from the average woman's, but also based on morphological cues to height and weight. Given that height and weight are sexually dimorphic, it is plausible that facial cues to these traits are used in forming perceptions of masculinity.

As the men in our sample were taller and had higher BMIs than women, this difference would have been reflected in the average male and female faces on which our masculinity scores were based. That individual variation in facial cues to height and BMI contributed to the perception of masculinity beyond the average shape dimorphism suggests an overgeneralization of facial trait correlates (see Section 2.1). Such overgeneralization is also revealed in the finding that both morphological scores of height and BMI were better predictors of perceived height and weight than were actual body height and BMI. Previous studies have shown that tall people have a more elongated face shape than short people (Mitteroecker et al., 2013; Windhager et al., 2011), and observers may overestimate differences in height based on this cue. Thus, the association between actual height and its facial correlates, such as elongation, may be overgeneralized to produce a perceptual relationship that is stronger than the correlation between facial correlates of height and actual physical height. Interestingly, we found that facial cues to height were negatively correlated with morphological masculinity, while both variables were positively linked to perceived masculinity. This may seem counterintuitive but it may be explained by focusing on two simplified characteristics associated with masculinity and height: width and elongation of faces. With increasing morphological masculinity, faces get wider; with increasing height, faces get longer. Both morphological masculinity and height are perceived as masculine (from a variety of surface traits, e.g., increased brow prominence), but the more elongated a face, the less wide it will be. This may partly account for the weak relationship between morphological and perceptual masculinity found in previous studies.

In order to interpret observed effects of masculinity on other interpersonal judgments, it is useful to understand the facial traits that influence perception of masculinity. The finding that the perception of facial masculinity is affected by not only sex-specific morphological features that are dependent on sex hormone levels, but also by traits that are linked to body size (independent of gender) aligns with studies on craniofacial allometry in humans and non-human primates (Mitteroecker et al., 2013; Schaefer, Mitteroecker, Gunz, Bernhard & Bookstein, 2004). Schaefer et al. (2004) suggested that the two dimensions of sexually dimorphic shape—sex-specific and size-dependent—may have been subject to different selection pressures; thus, they may have differential effects on social perceptions and preferences. That is, the effect that masculinity has on judgments such as attractiveness or leadership ability may depend not only on the extent to which a face is perceived to look masculine, but also on whether this perception of masculinity was formed based on cues to size or cues to sex hormone levels. Methods such as the one presented here provide the means to uncover distinct physical origins of social and stereotypic judgments that have to date been rolled into a singular concept of masculinity.

9. Study 2: Masculinity trade-off accounts revisited

Previous studies have argued that men's masculinity is linked in a curvilinear fashion to women's ratings of attractiveness, and that costs and benefits associated with choosing a masculine men may differ according to a range of aspects such as women's menstrual cycle status, environmental and individual condition. These factors will be reviewed in the following section. The presented experimental study will investigate attractiveness as a function of masculinity in more detail: While most previous studies investigated shifts in mean preferences, I investigated women's preferences for the same male faces at different levels of masculinity, establishing how tolerance for low and high masculinity levels might shift in different contexts.

9.1. Individual differences in women's preferences

Trivers (1972) theory on parental investment suggests that women have a higher minimal investment in offspring, and thus gain more in terms of reproductive success from the support of a potential partner than men. Based on this reasoning, it has been proposed that, in general, men should opt for short-term sexual encounters, whereas women should prefer long-term partner relationships. Moreover, men and women should show differential preferences for potential mates, responding to different adaptive problems they faced during evolutionary history. While both men and women have been found to value traits like intelligence and kindness as the most important characteristics in a partner, men place a special value on characteristics signalling high female reproductive potential, whereas women place higher value on cues to male parental investment, i.e., male capability and willingness to provide resources (e.g., Buss, 1989).

Nonetheless, it has been convincingly demonstrated that both men and women engage in preferential mate choice and same-sex competition over desirable mates—there is not only variation in mating strategies between the sexes, but equally important within-sex variation, caused by differential access to mates and environmental variation, and leading to shifts in preferences for certain mate traits (e.g., Burley, 1986; Buss & Schmitt, 1993; Gangestad & Simpson, 2000; Landolt, Lalumière & Quinsey, 1995). In order to be adaptive, mating strategies and preferences for mate characteristics such as apparent health, fertility, or willingness to invest in offspring can be expected to be facultative rather than absolute. The following section reviews evidence for such facultative preferences. Individual differences may account for much of the variance in observed preferences for faces in general, and sexual dimorphism in particular (e.g., DeBruine, Jones, Smith & Little, 2010).[1]

9.1.1. Menstrual cycle shifts

Gangestad and Thornhill (1998) were the first to observe a menstrual cycle shift in women's preferences for a male trait presumably linked to male quality: women in the fertile phase of their menstrual cycle were found to prefer the scent of more symmetrical men, while women in the non-fertile phase of their cycle did not seem to discriminate between the scent of less and more symmetrical men. Penton-Voak et al. (1999, see Section 8.1) found that a menstrual cycle shift could also be observed for women's preferences for male facial masculinity. These findings have been interpreted as evidence that women pursue a mixed-mating strategy whereby different male characteristics are preferred depending on whether a potential mate is assessed as a stable long-term partner or within the context of a short-term sexual encounter (e.g., Penton-Voak et al., 1999). Multiple studies have picked up this idea of a "hormone-mediated adaptive design" (Johnston et al., 2001) of preferences, and replicated that women seem to prefer more masculine faces during the fertile window of their cycle (Johnston et al., 2001; B. C. Jones et al., 2008, 2005; Little, Jones, Waitt et al., 2008; Penton-Voak & Perrett, 2000). Cyclic shifts in preferences have been found to also extend to non-facial traits such as voices, odour, and gait (briefly reviewed in Gangestad & Thornhill, 2008).

Recently, however, the existence of cycle shifts has been challenged which has sparked an ongoing discussion (DeBruine, Jones, Frederick et al., 2010; Gildersleeve et al., 2013; Harris, 2011, 2013; Harris, Chabot & Mickes, 2013), including two meta-analyses of overlapping samples that have come to opposing conclusions. While Gildersleeve, Haselton and Fales (2014a) argue that there is sound empirical evidence for cycle shifts in short-term relationship contexts (see Section 9.1.3), Wood, Kressel, Joshi and Louie (2014) conclude the evidence for cycle effects is not compelling (see also Gildersleeve, Haselton & Fales, 2014b; Harris, Pashler & Mickes, 2014; Wood, 2014; Wood & Carden, 2014, for further commentaries on discrepancies in findings and their interpretations).

A second line of criticism accepts the existence of cycle shifts, but challenges that cyclic preference shifts indeed represent an adaptation. Havlicek, Cobey, Barrett, Klapilova and Roberts (2015) argue that women's cyclical shifts in (masculinity) preferences do not reflect a

[1] Although a recent study claims that "context-effects" such as the ones reviewed in the following are of minor importance in explaining masculinity preferences compared to genetic differences (Zietsch, Lee, Sherlock & Jern, 2015).

mixed-mating strategy but are the by-product of an adaptive "individual sensitivity to ovarian hormone-dependent individual quality" (Havlicek et al., 2015, p. 8). That is, differences in, for example, estradiol levels between women would lead to differential mate value of women and related associated psychological adaptations. Cyclical hormone changes within women would lead to a shift in preferences as a by-product of this more general hormone-dependent mechanism that is based on between-individual differences.

9.1.2. Individual condition: attractiveness and health

Burley (1986) suggested that in species with biparental care, individuals have differential access to potential mates contingent upon their own desirability, which would also affect the amount of parental investment obtainable from a partner. Variation in individual condition may thus lead to different life history strategies.

How does this translate to facial preferences? Physical attractiveness, hypothesized to reflect "good condition", has been linked to differential preferences: female self-rated attractiveness is positively associated with preferences for male facial (and vocal) masculinity (Kandrik & DeBruine, 2012; Little, Burt, Penton-Voak & Perrett, 2001; O'Connor et al., 2012; Smith et al., 2009; Vukovic et al., 2008). Similarly, women who reported lower self-esteem were found to prefer less masculine faces (Johnston et al., 2001). This has been interpreted as evidence that "high quality females" may be able to acquire both good genes and investment from "high quality males", whereas for women with lower mate value the costs of selecting a low investing partner might be higher than the heritable health benefits that this partner might provide (Little et al., 2001). Subsequent studies found that the effect of individual condition interacts with the type of relationship women are looking for. Condition effects were only found in the context of long-term relationships (Little et al., 2001; Penton-Voak et al., 2003, see Section 9.1.3).

Little and Mannion (2006) found that exposing women to pictures of attractive same-sex faces decreased both their self-rated attractiveness and their preference for masculinity. They suggested condition-dependent mate choice should be better conceived as "market-value-dependent" mate choice: Masculinity preferences can be affected by manipulating self-rated opinion, and are thus affected by relative rather than absolute mate value. Interestingly, Buss (2008) suggested that it might be also for self-perceived mate value to mediate preference shifts during ovulation. Instead of demonstrating a shift in the value that is put on "good genes" markers *per se*, it may actually be women's self-perceived mate value that is shifting, "reflecting the fact that women actually become more reproductively valuable at ovulation" (Buss & Schmitt, 2011, p. 778; cf. Havlicek et al., 2015).

Individual condition and self-perceived mate value may also be linked to individual health.

Scott et al. (2008) found that self-reported health in a rural Malaysian population was positively associated with masculinity preferences, and significantly so only in a long-term but not short-term relationship context, a finding that mirrors previous findings on the effect of self-rated attractiveness discussed above. Nonetheless, two other studies suggest that health and perceived mate value have dissociable effects on mate preferences. Feinberg et al. (2012) found that while self-rated attractiveness and self-reported health were positively correlated, their effects on masculinity preferences differed. Self-rated attractiveness showed a positive association with masculinity preferences, but it was women with *poorer* self-reported health that preferred masculinity more in potential short-term partners than women who reported to be of better health. Feinberg et al. (2012) interpreted this finding as fitting with reasoning on pathogen disgust sensitivity (see Section 9.1.5.1): women who are more concerned about disease might place a higher value on cues to heritable health in a partner. In line with Feinberg et al. (2012), De Barra, DeBruine, Jones, Mahmud and Curtis (2013) predicted—and found—a negative correlation of childhood health and adult masculinity preferences, qualified by an interaction with present health. Individuals with currently poor health showed a stronger association of childhood diarrhoea and masculinity preferences than individuals with better health.

9.1.3. Relationship context: long- and short-term relationships

Little et al. (2001) suggested that the effect of own perceived mate-value on masculinity preferences might be moderated by the temporal relationship context in which men's attractiveness is assessed—i.e. whether women are judging men as potential short- or long-term partners. They reasoned that masculinity as a cue to parental investment should only be relevant in relationships that are expected to last for an extended period of time. Indeed, they found that self-rated attractiveness did not affect masculinity preferences in a short-term context, when facial cues to paternal investment are likely to be less important. If, however, asked to judge long-term attractiveness it was found that less attractive women preferred more feminine faces than more attractive women (cf. Section 9.1.2). Smith et al. (2009) found that a preference for less masculine men in a long-term relationship context was most pronounced in women who perceived less masculine men as particularly trustworthy.

In line with reasoning by Little et al. (2001), Penton-Voak et al. (2003) found that women rated by others as less attractive preferred more feminine faces in a short-term compared to a long-term context, while no such effect was observed for more attractive women. A general preference for more feminine faces in a long- as compared to a short-term context (independent of individual condition) was observed by Little, Jones, Penton-Voak, Burt and Perrett (2002), Penton-Voak et al. (2003), Scott et al. (2008) and Smith et al. (2009).

Additionally, masculine faces (and voices) have been found to be preferred over feminine ones for short-term relationships (Boothroyd, Jones, Burt, DeBruine & Perrett, 2008; Little, Connely, Feinberg, Jones & Roberts, 2011). Recently, it has also been suggested that ovulatory cycle shifts in preferences might be context-dependent, with shifts being more pronounced or only existent in short-term compared to long-term relationship contexts (Gildersleeve et al., 2014a; Little & Jones, 2012).

The importance of relationship type can be moderated by other factors, such as environmental condition. Little, Cohen, Jones and Belsky (2007) found that in harsh, insecure environments women prefer more feminine faces in a long-term than in a short-term relationship context. If, however, primed with cues to a safe environment, no effect of relationship type was observed (see Section 9.1.5.2).

9.1.4. Relationship status

Little et al. (2002) suggested that masculinity preferences should be higher in partnered compared to single women: in a stable, happy relationship, women might assess other men's attractiveness in the context of a short-term sexual encounter, where facial cues to pro-sociality are less likely to be important or possible good-gene benefits might be maximized in an extra-pair copulation. In line with this reasoning, Little et al. (2002) and Sacco, Jones, DeBruine and Hugenberg (2012) found that happily partnered women preferred more masculine faces than single women.

9.1.5. Environmental conditions

9.1.5.1. Pathogen load

If there is a trade-off between genetic quality and paternal investment in potential mates, and if male genotypic quality is indicated by facial masculinity (see earlier), women should prefer more masculine faces when there is high pathogen load in the environment. In general, populations with high pathogen load place more value on physical attractiveness in mate choice (Gangestad & Buss, 1993; Low, 1990).

Penton-Voak, Jacobson and Trivers (2004) compared female masculinity preferences in Jamaica and the UK, two countries that differ in parasite load and health care. They demonstrated a stronger preference for male masculinity in women from Jamaica than in women from the UK. Two interpretations are possible: either health as reflected by male facial masculinity is indeed higher valued in Jamaica; or there is a populational difference in how much importance in mate choice is attached to cues of more positive personality characteristics (Penton-Voak et al., 2004).

DeBruine, Jones, Crawford, Welling and Little (2010) investigated cross-cultural masculinity preferences with regards to health in 30 countries. They found average population preferences for masculinity to increase with decreasing national health as calculated from eight WHO statistics ($r=-.62$). Brooks et al. (2011) suggested this relationship to be mediated by income inequality, which would not only predict health, but also prevalence of (violent) male-male competition. This was rebutted by DeBruine et al. (2011) by controlling for gross national product in their 2010 dataset, and collecting new data from over 8000 women in the US—national health in different US states was a better predictor of masculinity preferences than income equality or homicide rate.

In addition to findings on nationwide variation, preferences have been also found to be contingent on personal differences in sensitivity to pathogens. Exposing raters to pictorial cues of pathogen contagion was found to shift men and women's preferences towards more sexually dimorphic opposite-sex faces (Little, DeBruine & Jones, 2011), and women's preference for masculine male faces is positively correlated with self-reported pathogen disgust (DeBruine, Jones, Tybur, Lieberman & Griskevicius, 2010; B. C. Jones et al., 2013). In line with reasoning that masculinity reflects genetic fitness, Lee and Zietsch (2011) found that when primed with cues of pathogen prevalence, women place higher values on "good-gene" traits than "good-dad" traits. Recently, however, the link between pathogen disgust sensitivity and masculinity preferences has been called into question. Lee and Zietsch (2015) found that pathogen disgust sensitivity only affected masculinity preferences if measured with forced-choice paradigms, and effects seemed to be limited to women of relatively young age (i.e. less than 35 years old).

9.1.5.2. Resource availability

Little, Cohen et al. (2007) found that when women were primed with verbal cues to an environment scarce of financial (but also emotional) resources, women preferred more feminine faces than when primed with cues to a safe environment. This effect was only visible in a long-term but not a short-term relationship context. They concluded that harsh environments may favour the choice of lower-quality but higher-investing long-term partners.

Little, DeBruine and Jones (2013) used a similar priming paradigm but included a control condition—women indicated their masculinity preference prior to experimental manipulation, and after being primed with visual cues to low and high wealth. This allowed a more refined analysis of preferences: not only did masculinity preferences decrease compared to baseline when primed with cues to low wealth, but it was also found that masculinity preferences increased when women were primed with cues to high wealth, suggesting that women prefer more feminine men in resource-low, but more masculine men in resource-high environments. In line with these findings, Lee and Zietsch (2011) found that women more generally prefer "good-dad" traits over "good-gene" traits when primed with resource scarcity.

A similar reasoning had already been put forward by Gangestad and Buss (1993): female access to resources should change the trade-off women face between mates that provide high investment and mates that have higher genetic fitness. Higher (financial) status should lead women to place more importance on mate characteristics relevant to fitness relative to characteristics relevant to exclusive investment. Moreover, all else being equal, having greater access to resources leads to an increase in women's fitness and hence should affect women's (perceived) mate value. Indeed, Moore, Cassidy, Law Smith and Perrett (2006) showed that women's resource control was associated with a greater preference for physical attractiveness compared to good financial prospects in a partner. Resource availability might thus be linked to facial masculinity preferences through an effect on female condition.

9.1.6. Personality, attitudes and experience

Individual differences in opposite-sex facial preferences have also been related to factors such as personality and individual experience. The positive relationship between women's self-rated attractiveness and their preference for masculine men's faces was found to be mediated by extraversion (Welling, DeBruine, Little & Jones, 2009), and women's masculinity preferences were found to be positively related to a "feminine" personality trait, i.e. empathising, while men's femininity preferences were linked to a "masculine" trait, systemizing (Smith, Jones & DeBruine, 2010). Welling, Jones and DeBruine (2008) reported a positive association between women's masculinity preferences and their self-rated sex drive. Waynforth et al. (2005) linked masculinity preferences to female mating strategies using the sociosexual orientation inventory (SOI). A high score on this measure reflects higher self-reported sexual experience and interest in short-term relationships. Although women were found to slightly prefer overall more feminine male faces, higher scores on the sociosexual orientation inventory were associated with a greater preference for masculine male faces. Similarly, Boothroyd and Brewer (2014) found a positive correlation of overall SOI scores and masculinity preferences (but no such effect for the SOI attitude sub scale). However, this relationship became non-signifcant when entering SOI scores into a regression model together with measures of impulsivity—planning behavior (or lack thereof) was the only significant predictor of masculinity preferences (Boothroyd & Brewer, 2014).

It has been also suggested that preferences for facial sexual dimorphism can be affected by visual diet. Saxton, Little, DeBruine, Jones and Roberts (2009) found that female adolescents who attended a same-sex schools showed a lesser preference for masculinity in opposite-sex faces than adolescents who attended a mixed-sex school (or had opposite-sex siblings). Boothroyd and Perrett (2008) found that early family stress, too, can impact on masculinity preferences: the quality of women's relationships with their parents was positively linked to masculinity preferences.

9.1.7. Study rationale and predictions

Previous studies on the effect of individual differences on masculinity preferences were based on two main premises. (1) Attractiveness and masculinity are related in a curvilinear fashion. Very low as well as exceedingly high levels of masculinity are unattractive. The former for indicating low levels of testosterone, and hence a "low-quality" mate; the latter for being associated with undesirable social traits, such as unfaithfulness, coldness, and a lack of trustworthiness. (2) Preferences for masculinity are subject to a trade-off that is affected by several factors, such as hormonal and environmental conditions.

To our knowledge, the first assumption has never been explicitly tested.[2] The first aim of the current study thus was to investigate whether the relationship between masculinity and attractiveness is indeed curvilinear. The second assumption has been derived from findings that show that mean preferences for masculinity can differ between, and shift within, women depending on the factors discussed above. It was concluded that costs and benefits associated with choosing a masculine partner differ depending on these aspects. Yet, as B. C. Jones et al. (2013) noted, the use of experimental designs such as a two-alternative forced choice task cannot answer the question whether individual differences in relative preferences for masculinised vs feminised faces are driven by an increased attraction to masculine men and/or an increased aversion to feminine men. Here, we directly investigated potential differences in attraction to low and high levels of masculinity. We determined the attractiveness of the same men at different levels of masculinity, thereby investigating the tuning of facial preferences, i.e. rated attractiveness, as a function of masculinity in facial shape.

Most previous studies investigating the effect of male facial masculinity on women's preferences have worked with 2D photographs. Often, masculinity of individual faces was manipulated and presented to participants either in two-alternative forced choice tasks or in interactive tasks asking users to maximise attractiveness by hovering the mouse over a stimulus face, thereby changing its masculinity. The current study introduced two methodological changes; the use of 3D stimulus faces (1), manipulated in masculinity but presented in a rating task (2). Thus, I also tested whether there might be systematic differences in masculinity preferences as assessed with 2D and 3D stimuli, and whether previously established effects of individual differences on masculinity preferences replicated in the current sample when

[2] Note that while Said and Todorov (2011) modelled a quadratic relationship of face shape components and attractiveness, they did not explicitly test for a quadratic relationship of masculine face shape and attractiveness but compared the direction of the vector of greatest attractiveness to the direction of the vector of sexual dimorphism.

tested with a previously used method, i.e. an interactive task. The following hypotheses were tested.

Self-rated attractiveness Based on the finding that less attractive women prefer more feminine male faces than highly attractive women, and prefer more feminine faces in a long- compared to a short-term relationship context, the following prediction was tested.

> *A.* Women rating themselves as more attractive should show a reduced tolerance of low masculine male faces and/or a higher tolerance of highly masculine male faces than women perceiving themselves as less attractive, and especially so when comparing preferences in a long-term and short-term relationship context.

Relationship status It has been suggested that women who are in stable long-term relationships would assess other men's attractiveness in the context of a potential extra-pair copulation, for which a) facial cues to parental investment (i.e. low masculinity) are less important, and b) facial cues to good health (i.e. high masculinity) are more important. Thus, the following prediction was tested.

> *B.* Women who have a partner (and are happy in/committed to their relationship) should show a higher preference for/a higher tolerance towards high levels of masculinity than women who are currently single.

Environmental harshness: financial insecurity Previous studies suggest that priming women with cues to financial scarcity/environmental harshness decreases their masculinity preferences. Instead of priming participants with hypothetical scenarios, we tested whether perceived financial resource scarcity affects masculinity preferences: We asked women how much they worried about their future financial situation, and tested the following prediction.

> *C.* Women who worry more about their financial future show a higher tolerance/ preference for low levels of masculinity than women who are less worried about their financial future.

Pathogen disgust sensitivity Facial masculinity has been suggested to be linked to heritable good health. As DeBruine, Jones, Tybur et al. (2010) suggested, health benefits might be off-setting costs of high masculinity when pathogens are a greater concern. We thus tested the following prediction.

> *D.* Women with higher pathogen disgust sensitivity show a higher preference/ tolerance for high levels of masculinity than women scoring low on pathogen disgust sensitivity.

Self-reported health Based on the presumed link of masculinity and heritable health benefits, it could be predicted that for women with poor health, partner's health is of greater value than for women of self-reported good health. However, health might affect self-perceived mate value, and a prediction of the opposite direction is conceivable—that for women with poor health, and therefore lower mate value, the costs of choosing a highly masculine (and low-investing mate) outweigh benefits, reflected in a preference for feminine men. If self-reported health is indeed linked to (self-perceived) mate value, preferences might also differ for long- and short-term relationship contexts.

E. Self-reported health affects masculinity preferences.

Control variables **Age** Little et al. (2010) found that masculinity preferences were highest for women within a reproductive age range. Moreover, several studies have indicated a positive correlation of age and masculinity preferences within a reproductive age range (e.g., Little et al., 2001, 2002, but see, for example, DeBruine et al., 2006, for a null-finding regarding age). Analyses were thus limited to women aged 18–45, and age was controlled for in all analyses.

Sexual orientation Batres, Jones, Feinberg, DeBruine and Perrett (n.d.) found that variation in women's attraction to men predicted masculinity preferences. They asked participants to indicate their sexual orientation on a variant of the Kinsey scale (Kinsey, Pomeroy & Martin, 1948), a continuous scale ranging from 1-*homosexual* over 4-*bisexual* to 7-*heterosexual* (e.g., Boothroyd et al., 2008; Moore et al., 2006). For women who identified their sexual orientation as 5, 6 or 7 on this scale, a positive relationship of sexual attraction to men and masculinity preference was observed. A preliminary base model tested for the effect of sexual orientation on masculinity point preference measures and tuning curves before deciding on how this variable would be treated in subsequent analyses of individual differences (see Section 9.3.1).

Hormonal contraceptive use Hormonal contraceptives appear to cause systematic differences in masculinity preferences. Little, Burriss, Petrie, Jones and Roberts (2013) showed that initiating hormonal contraceptive use led to a decrease in women's preference for male facial masculinity. They also found that the partners of women who were using hormonal contraceptives when the relationship was formed had less masculine faces than the partners of women who were not using hormonal contraceptives at that time, suggesting not only a difference in mate preferences but indeed mate choice. Menstrual cycle shifts in preferences have been found to be weaker, or non-existent in women who are using hormonal contraceptives (Alvergne & Lummaa, 2010; Penton-Voak et al., 1999). Hormonal contraceptive use can also moderate the effect of factors such as self-rated attractiveness (see Section 9.1.2), relationship context (see Section 9.1.3) and relationship status (see Section 9.1.4). Little et al. (2002) found that relationship context and being single or not only affected women's masculinity preferences when they were not using hormonal contraceptives. Similarly, Smith et al. (2009) found that only masculinity preferences of women who were not using the pill were affected by the temporal context of relationships. As Little et al. (2002) noted, differences in preferences between women using and not using hormonal contraceptives might also be caused by behavioural differences—women who are using hormonal contraceptives might also engage in behaviours or make lifestyle-choices that are different from those of women not using hormonal contraceptives (e.g., women on the pill reported more sexual partners than women not using the pill).

As with sexual orientation, the effect of hormonal contraceptive use on masculinity preferences was first tested in a preliminary base model before deciding on how this variable would be treated in further analyses (see Section 9.3.1).

9.2. Methods

9.2.1. Participants

A total of 563 women were recruited through the Perception Lab website (*online* sample) and through Amazon MTurk (*MTurk* sample, Buhrmester, Kwang & Gosling, 2011). Amazon MTurk workers were paid $2.00 for their participation. Exclusions were made based on ethnicity (Caucasian only), age (only women in a reproductive age range, i.e. between ages 18 and 45), sexual orientation (only women who reported higher sexual attraction to men than women) and rating behaviour (only women who assigned more than two different values when judging men's attractiveness on an 8-point Likert-type scale, see Section 9.2.3). Table 9.1 provides descriptive statistics for the overall and final sample.

Table 9.1.: Descriptive statistics for Study 2.

		Total		After exclusions	
		Online	MTurk	Online	MTurk
N		267	296	173	142
Age	*M*	24.98	37.08	24.12	33.51
	SD	8.84	11.53	7.26	6.27

9.2.2. Stimulus set

The stimulus set consisted of four base faces that were composites of four men each (see Figure 9.2.1). Men were picked from Face Set 2 (see Section 4.2) to match in age, height, and BMI (see Table 9.2).

Table 9.2.: The composite base faces used in Study 2 were matched in age, height, and BMI. The table gives the range in mean age, height and BMI of the four composite face in the 3D and 2D tasks, as well as the range for the individual faces that went into the respective composites.

		Age (years)	Height (cm)	BMI (kg m-2)
3D stimuli	Composite base faces	20.8 - 22.0	180.1 - 182.5	22.0 - 22.4
	Individual faces	18.0 - 25.0	173.0 - 191.5	18.0 - 25.0
2D stimuli	Composite base faces	21.0 - 22.5	180.3 - 180.8	21.9 - 22.4
	Individual faces	19.0 - 26.0	174.0 - 191.0	19.0 - 26.0

Figure 9.2.1.: Base faces for Study 2. Each base face was a composite of four men from Face Set 2.

Base faces were manipulated in their masculinity by applying or subtracting the linear difference between the average male and female face shape from Face Set 2. With this difference corresponding to 100%, each base face was feminized and masculinised to cover a range of –100% to +200% sexually dimorphic shape in seven steps of 50%, resulting in 28 stimulus faces. Figure 9.2.2 shows one of the base faces at the seven different levels of masculinity. Note that transforming faces in this way changes face shape along the male-female axis while retaining identity and sex. The (asymmetric) range of –100% to +200% was chosen based on a pilot study, which showed that when presented with a range of –100% to +100%, women rated very low levels of masculinity as unattractive, but very high levels of masculinity as relatively attractive, indicating a ceiling effect.

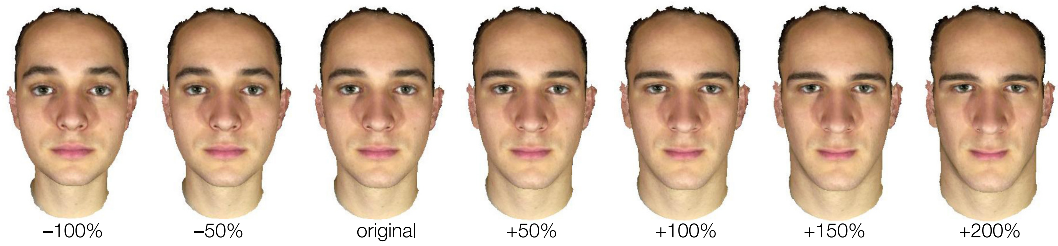


Figure 9.2.2.: One of the base faces at the seven different levels of masculinity. Masculinity transforms were based on the difference in average face shape of men and women from Face Set 2.

9.2.3. Experimental tasks

9.2.3.1. Task 1: 3D rating

The stimulus faces were presented in two separate tasks. In Task 1, women were asked to rate all 28 faces on their attractiveness on an 8-point Likert-type scale from 1–*Not at all attractive* to 8–*Very attractive*. Prior to the rating, participants were presented with 2D frontal images of all face models for one second each to provide an overview of stimulus variability. The 3D face stimuli were presented on a computer screen in randomized order. They were rotated from –45° to +45° from left to right while simultaneously being rotated from –15° to +15° up and down, resulting in the stimuli "bobbing" in a sinusoidal manner. Images were presented individually against a black background and remained visible until a rating was made.

9.2.3.2. Task 2: 3D and 2D interactives

In Task 2, the stimulus faces from Task 1 were presented in an interactive slider task that showed each face from a frontal and lateral view (see Figure 9.2.3)[3]. Interactive stimuli were presented in two randomised blocks. In each block, participants were briefed with one of the following two instructions (Penton-Voak et al., 2003):

> (1) "In this block, you are asked to imagine you are looking for a SHORT-TERM partner. Short-term implies that the relationship may not last a long time. Examples of this type of relationship would include a single date accepted on the spur of the moment, an affair within a long-term relationship, and possibility of a one-night stand."

> (2) "In this block, you are asked to imagine you are looking for a LONG-TERM partner. Examples of this type of relationship would include someone you may want to move in with, someone you may consider leaving a current partner to be with, and someone you may, at some point, wish to marry (or enter into a relationship on similar grounds as marriage)."

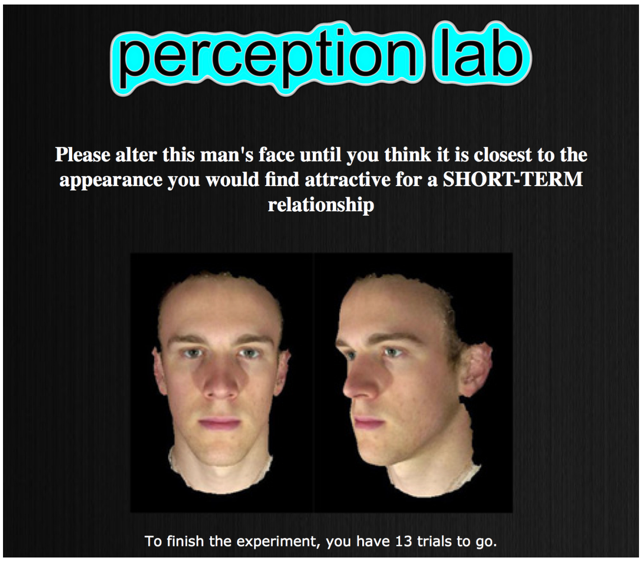


Figure 9.2.3.: Interface of Task 2. By hovering the mouse over the face, participants changed its masculinity simultaneously in a frontal and lateral view.

To compare findings to previous studies using 2D stimuli, Task 2 also included a 2D interactive task. Again, four composite faces of four men each served as base identities that

[3]A demo of the interactive interface can be found at `http://perceptionlab.com/~Iris/SI.html`.

were transformed to cover a masculinity range from–100% to +200% (see Figure 9.2.4). Faces were chosen from a database of commercially available facial pictures (3DSK, e.g., Re et al., 2013) to match the 3D stimulus faces in age, height and BMI (see Table 9.2). Masculinity transforms were based on the average of 50 women and 50 men from the same data base.

Figure 9.2.4.: Stimulus faces in 2D control condition. Each face was a composite of four men each from the 3DSK picture set.

The order of experimental tasks was fixed; participants first completed Task 1 (3D rating), then Task 2. The order of blocks within Task 2 (short-term, long-term) was randomised; within these blocks, participants first completed the 3D, then the 2D interactives; the order of faces within each interactive task was randomised.

9.2.4. Instructional manipulation check

For the *MTurk* sample, an instructional manipulation check (IMC) was included prior to the experimental tasks to ensure participants carefully read instructions. The IMC described in Oppenheimer, Meyvis and Davidenko (2009) was adapted to the current experiment. Participants were presented with the following instruction:

> This experiment investigates individual differences in preferences for face shape and skin tone. Individual preferences and knowledge, along with situational variables, can greatly impact decision processes. In order to facilitate our research we are interested in knowing certain factors about you, the decision maker. Specifically, we are interested in whether you actually take time to read the instructions; if not, some of our tests that rely on changes in the instructions will be ineffective. So, in order to demonstrate that you have read the instructions, please ignore the question below, as well as the "Next" button. Instead, simply click on the "Back" button to proceed to the next screen. Thank you very much.

Beneath this paragraph, participants were presented with the question "Which of these factors influence facial attractiveness the most in your opinion?" and could rate the importance of aspects such as "clear skin" or "a beautiful smile". At the very bottom, participants could either click "Back" or "Next". If they clicked "Back" the experiment proceeded; if they clicked "Next", they were taken to a page that explained the importance of reading instructions and told participants they could restart the experiment if they wanted to.

9.2.5. Individual differences

Prior to the experimental tasks, participants were asked to fill out a questionnaire on basic demographic information (age, gender, ethnicity).

Participants indicated their sexual orientation on a 7-point scale ranging from 1-*homosexual* to 4-*bisexual* to 7-*heterosexual* (Moore et al., 2006). Only participants that reported a sexual orientation of 5 or above (i.e. a greater sexual attraction to men than women) were included in the subsequent analyses. Health was measured on a 5-point scale with the options 1-*Excellent*, 2-*Very good*, 3-*Good*, 4-*Fair* and 5-*Poor* (e.g., Jürges, Avendano & MacKenbach, 2008). Answers on the health item were reverse-coded for analysis, so that higher values corresponded to better health. Disgust sensitivity was measured with the seven-item pathogen disgust sensitivity scale (Tybur, Lieberman & Griskevicius, 2009). Items were summed to give a disgust sensitivity score, with high values indicating a high disgust sensitivity. If a participant did not rate all seven items, their disgust sensitivity was recorded as missing. Participants rated their own attractiveness (to the sex they were attracted to) on a 7-point Likert-type scale ranging from 1-*Below average/Not so attractive* to 7-*Above average/Very attractive*. Participants were also asked about their relationship status: whether they were currently in a relationship, and if so how happy they were in the relationship (7-point Likert-type scale ranging from 1-*Very unhappy* to 7-*Very happy*), and how committed they felt to their relationship (7-point Likert-type scale ranging from 1-*Not committed at all* to 7-*Very committed*). Only women who reported to be committed to and happy in their current relationship (values of 4 or above on the respective scales) were included in the analysis of the effect of relationship status (Little et al., 2002). To approximate perceived environmental harshness, participants were asked to report how much they worried about their future financial situation when thinking ahead (7-point Likert-type scale ranging from 1-*I worry a lot* to 7-*I'm not worried at all*). Finally, participants were asked whether they were currently using hormonal contraceptives, and whether they had answered all questions truthfully (all women in the final sample—i.e. after the described exclusions—indicated they had).

9.2.6. Analyses

As few studies have worked with 3D faces so far, masculinity preferences were first compared across the different modalities (2D and 3D stimuli) and tasks (interactive tasks vs. rating task). Next, we tested for the effects of the recorded individual differences on masculinity preferences.

(1) Preferences across different modalities and tasks. To compare preferences in 2D and 3D faces, preferred masculinity levels in the interactive tasks were averaged across the four base faces as well as long-term and short-term contexts for each participant, separately for 2D and 3D faces. To allow for a comparison of preferred masculinity levels in the interactive tasks and the rating task, each woman's peak preference level in the 3D rating task was computed. Peak preference was determined as the level to which the highest attractiveness rating had been assigned. For each woman, the mean attractiveness rating for each masculinity level was computed by aggregating ratings across the four base faces. If the same rating was given to two consecutive masculinity levels, the peak preference level was calculated as the average of those two levels (e.g., if the highest rating was given to levels +50% and +100%, peak preference was determined as 75%). If the same rating was given to more than two levels, or two levels that were not consecutive, peak preference was recorded as missing. Correlation analysis was used to test for an association of masculinity preferences across different modalities and tasks.

(2) The effect of individual differences on masculinity preferences in interactive and rating tasks. As the data had a hierarchical structure (four base faces assessed by participants from two different samples), linear mixed effect models were used to test for the effect of both control variables as well as individual differences in the variables of interest. All tested models are presented in the following form: *outcome variable ~ (predicted by) fixed effects + random effects*.

For preferences in the interactive tasks, the initial model included the chosen level of masculinity as outcome variable and two fixed effects, stimulus dimension (2D vs 3D) and context (short-term vs long-term). Random effects due to differences between the *online* and *MTurk* samples, differences in participant scale use and differences in the baseline attractiveness of the four base faces were modelled by including a random intercept for faces nested in participant nested in sample. After establishing the effect of the three control variables, this base model was then adapted for the analyses of the different variables: questionnaire items were entered as additional fixed effects, and, depending on the specific research question, allowed to interact with dimension and/or context.

To analyse attractiveness as a function of masculinity, attractiveness ratings were entered as the dependent variable, and masculinity level as a predictor. As we expected the relationship of attractiveness and masculinity to be curvilinear, we added a quadratic term of masculinity level to the model. Random effects were modelled as for the interactive task. For the analysis of the different variables, questionnaire items were then entered as fixed effects and allowed to interact with both the linear and quadratic masculinity level term.

All analyses were carried out using *SPSS 22* and *R* (R Development Core Team, 2015). Linear mixed effect models were calculated using the *R* packages *lme4* (Bates, Maechler, Bolker & Walker, 2015) and *lmerTest* (Kuznetsova, Brockhoff & Christensen, 2015). Interaction plots were created using *sjPlot* (Lüdecke, 2015); interactions with continuous variables were by default plotted for the marginal means for the minimum and maximum values of the independent variable. All *p*-values are reported two-tailed.

9.3. Results and discussion

9.3.1. Preliminary analyses

9.3.1.1. Masculinity preferences across different modalities and task

Masculinity preferences expressed in the 3D and 2D interactive tasks were averaged across base faces and relationship contexts. Mean preferences for 3D and 2D faces were strongly correlated ($r(313)=.57$, $p<.001$, see Figure 9.3.1). Peak preference as measured in the rating task was found to be correlated with preferences in both 3D ($r(279)=.65$, $p<.001$) and 2D ($r(279)=.41$, $p<.001$) interactive tasks (see Figures 9.3.2 and 9.3.3).

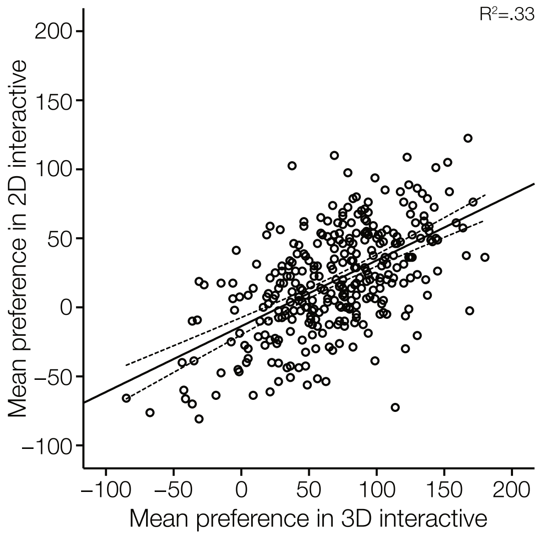


Figure 9.3.1.: Masculinity preferences were highly correlated for 3D and 2D interactive tasks. The solid line is the line of best fit; dashed lines show the 95% CI.

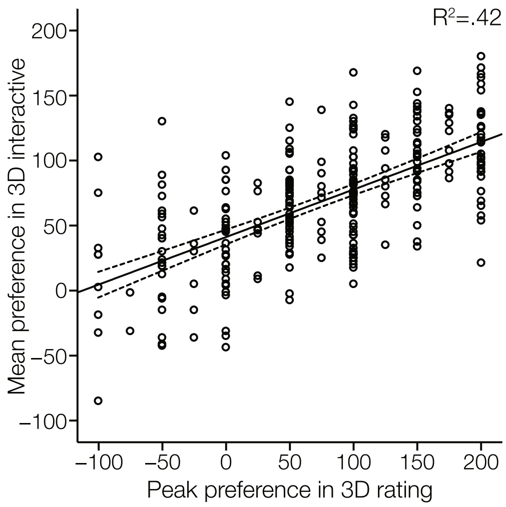


Figure 9.3.2.: The masculinity level chosen to be the most attractive in the rating task correlated with mean masculinity preferences in the 3D interactive task. The solid line is the line of best fit; dashed lines show the 95% CI.

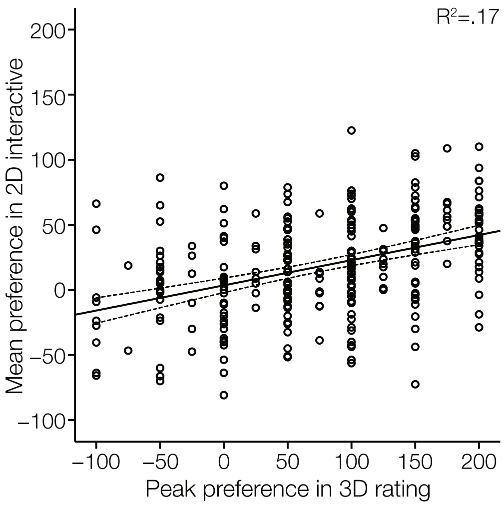


Figure 9.3.3.: Peak preference level in the rating task was also correlated with preferences in the 2D interactive task, albeit less strongly so. The solid line is the line of best fit; dashed lines show the 95% CI.

While preferences in the different tasks were correlated, mean preferences in the three tasks differed. The highest preference was expressed in the 3D rating (M=76%), followed by the 3D interactive (M=69%) and the 2D interactive (M=19%; see Figure 9.3.4). One-sample t-tests showed that preferences in all three tasks were significantly higher than the original (i.e. unmanipulated) level of masculinity (3D rating: $t(279)=15.60$, $p \leq .001$; 3D interactive: $t(313)=26.34$, $p \leq .001$; 2D interactive: $t(313)=8.86$, $p \leq .001$).

9.3.1.2. Analysing preferences in interactives and rating: base models

Before testing individual effects, I determined effects of the three control variables: participant age, sexual orientation and hormonal contraceptive use.

Interactive tasks I first tested a linear mixed effect model with chosen masculinity level as outcome variable, the predictors stimulus dimension (2D vs 3D) and context (short-term vs long-term) and a random intercept for stimulus face nested in participants nested in sample. Then, participant age, sexual orientation and hormonal contraceptive use were added to the model. To test whether effects might differ for 2D and 3D stimuli, all three control variables were allowed to interact with stimulus dimension. The model including age, sexual orientation and hormonal contraceptive use was found to be a better fit (AIC=14082) than the model without these control variables (AIC=14104, $\chi^2=34.33$, $p<.001$).

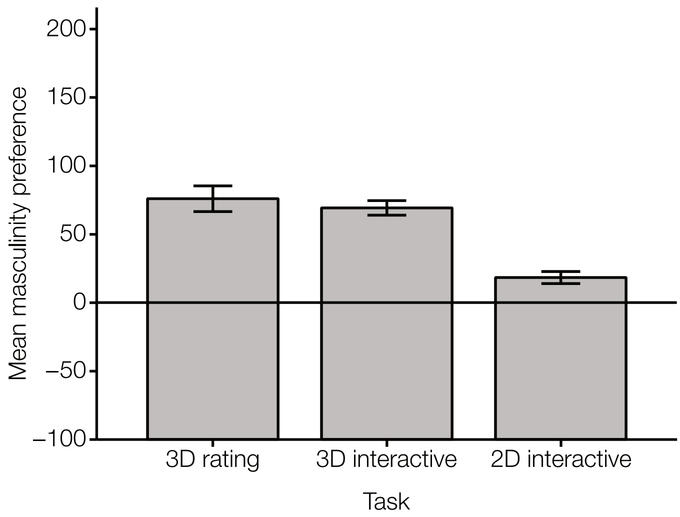


Figure 9.3.4.: Masculinity preferences in the different tasks. The solid line marks the masculinity level of unmanipulated faces. In all three tasks, preferences were significantly different from 0, indicating an overall preference for increased masculinity. This preference was more pronounced in 3D compared to 2D tasks.

Results showed significant main effects of context (short-/long-term) and dimension (2D/3D) on masculinity preferences (see Table 9.3). Preferences for masculine faces were higher in a short- compared to a long-term context, and greater for 3D compared to 2D faces (cf. Section 9.3.1.1). These main effects were qualified by an interaction between relationship context and stimulus dimension, whereby higher masculinity was preferred in a short-term compared to long-term context to a greater extent in 3D compared to 2D faces (see Figure 9.3.5).

The main effect of participant age on masculinity preferences was not significant but participant age showed a significant interaction with stimulus dimension: older women showed a greater preference for higher levels of masculinity than younger women in 2D but not 3D faces (see Figure 9.3.6). Sexual orientation had a significant main effect on masculinity preferences: the higher women's score on the sexual orientation scale (i.e., the more women reported to be exclusively attracted to men), the higher their preferred level of masculinity. Sexual orientation did not interact with stimulus dimension in predicting masculinity preferences. Hormonal contraceptive use showed no significant main effect on preferences, but interacted with stimulus dimension: women using hormonal contraceptives preferred a higher level of masculinity than women not using hormonal contraceptives in 2D compared to 3D faces.

Table 9.3.: Masculinity preferences in interactive task—base model including the control variables participant age, sexual orientation and hormonal contraceptive use. Estimates for categorical variables are given with respect to reference categories; these were "context *short-term*", "dimension *3D*" and "contraceptive use *no*".

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	2.59	0.47	376	5.57	<.001
context	−0.17	0.04	4626	−4.59	<.001
dimension *2D*	−0.82	0.29	4626	−2.83	.005
dimension *2D* x context *long-term*	0.11	0.05	4626	2.03	.042
age	0.01	0.01	374	1.36	.174
age x dimension *2D*	0.01	0.00	4626	2.61	.009
sexual orientation	0.24	0.07	374	3.34	.001
sexual orientation x dimension *2D*	−0.08	0.05	4626	−1.68	.093
contraceptive use *yes*	0.08	0.10	374	0.79	.429
contraceptive use *yes* x dimension *2D*	0.16	0.06	4626	2.71	.007

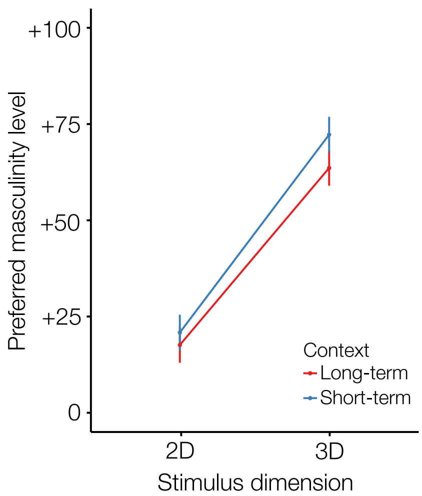


Figure 9.3.5.: Women preferred a higher level of masculinity when optimising men's facial attractiveness in a short-term compared to a long-term relationship context, and especially so for 3D faces. Error bars represent 95% CI.

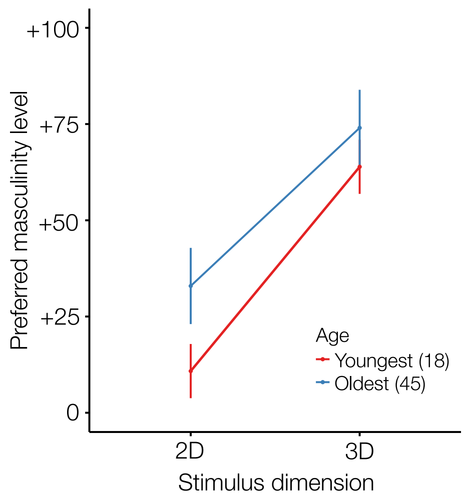


Figure 9.3.6.: Older women preferred a higher level of masculinity than younger women, and especially so in 2D compared to 3D faces. Error bars represent 95% CI.

Tuning curves I first predicted ratings of attractiveness by entering men's level of facial masculinity as a linear term. Re-running the model including the quadratic term significantly increased model fit (AIC 30240 vs 29499, χ^2=743.4, p<.001). In a next step, the three control variables (participant age, sexual orientation and hormonal contraceptive use) were added to the model and allowed to interact with both linear and quadratic masculinity terms. Model fit significantly improved by adding the control variables (AIC=26379, χ^2=137.6, p<.001).

Table 9.4 summarizes the results. Age had a significant main effect and was found to interact with both the linear and the quadratic masculinity term. Hormonal contraceptive use had no main effect on attractiveness ratings, nor did it interact with either the linear or quadratic masculinity terms. Sexual orientation had a main effect on attractiveness ratings, and interacted with the linear but not the quadratic masculinity term.

Table 9.4.: Masculinity preferences in rating task—base model of preference curves including the control variables participant age, sexual orientation and hormonal contraceptive use. The estimate for contraceptive use is given with respect to the reference category, "contraceptive use *no*". For brevity, masculinity level is from here on abbreviated as "mlevel", and the quadratic term as "mlevelsq".

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	4.86	0.84	179	5.79	<.001
mlevel	0.17	0.32	7480	0.53	.597
mlevelsq	−0.09	0.04	7480	−2.16	.031
age	0.02	0.01	747	2.03	.043
age x mlevel	−0.01	0.00	7480	−2.69	.007
age x mlevelsq	0.00	0.00	7480	2.45	.014
contraceptive use *yes*	−0.20	0.16	1079	−1.21	.227
contraceptive use *yes* x mlevel	0.08	0.07	7480	1.25	.210
contraceptive use *yes* x mlevelsq	−0.00	0.01	7480	−0.03	.978
sexual orientation	−0.52	0.12	1068	−4.14	<.001
sexual orientation x mlevel	0.16	0.05	7480	3.18	.001
sexual orientation x mlevelsq	−0.01	0.01	7480	−1.22	.224

To visualize the shape of the preference function, a curve was fitted based on the estimated slopes from the linear mixed effect model. The local maximum of the function (for values corresponding to the range of presented masculinity levels, 1–7) was calculated as a measure of peak preference and graphically illustrated. Figure 9.3.7 visualizes masculinity preference curves at the sample minimum and maximum age (18 and 45 years, respectively) for women who were not using hormonal contraceptives and reported to be completely heterosexual. Figure 9.3.8 visualizes masculinity preference curves at the sample minimum and maximum sexual attraction to men (5 and 7 on the sexual orientation scale, respectively) at the sample mean age (28.4 years) when no hormonal contraceptives were used.

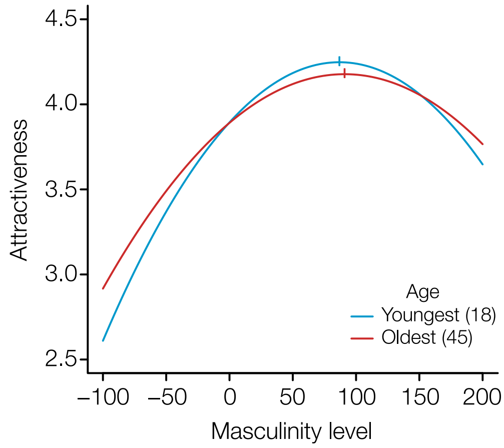


Figure 9.3.7.: Attractiveness as a function of masculinity in younger and older women. The ticks mark the turning points of the preference curves or peak preference levels. The peak preference level for an average 18-year old woman in the sample was +87% masculinity, while the peak preference for an average 45-year old was +91% masculinity.

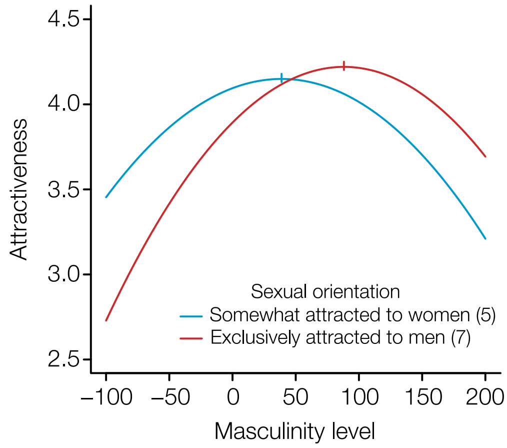


Figure 9.3.8.: Attractiveness as a function of masculinity in women with varying sexual attraction to men. The peak preference level for (heterosexual) women reporting the strongest attraction to women was +39% masculinity, while the peak preference for women who reported to be exclusively attracted to men was +88% masculinity.

9.3.1.3. Discussion

Masculinity preferences were found to be strongly correlated in the three different tasks (2D interactive, 3D interactive, 3D rating). Moreover, masculinity preferences in all three tasks were significantly different from zero. This is in line with findings from previous studies that have reported a general preference for masculinity/masculine traits (Cunningham, Barbee & Pike, 1990; Gillen, 1981; Grammer & Thornhill, 1994; Koehler et al., 2004; Neave, Laing, Fink & Manning, 2003; Rhodes et al., 2003, 2007; Saxton et al., 2009; Scheib et al., 1999), but in contrast to other studies, most of which used a similar methodology as the current study and reported that, overall, women prefer a close to average or slightly feminine male face shape (Little et al., 2001; Penton-Voak et al., 2004, 2003; Perrett et al., 1998; Rhodes et al., 2000; Scott et al., 2010). Our results might differ from these latter findings because we used an asymmetric range of masculinity (−100% to +200%). Being presented with more masculinised compared to feminized faces, participants might have shifted their preferences towards a higher level of masculinity/the average of the presented range (+50%).

My second main finding was that while masculinity preferences were correlated in 2D and 3D, they were significantly greater in 3D compared to 2D faces. As 2D faces were presented with the same asymmetric masculinity range in masculinity, range effects cannot account for systematic differences between 2D and 3D faces.

Two explanations for systematic differences between 2D and 3D faces seem possible. First, masculinity manipulations might change different traits/traits differently in 2D and 3D faces. For example, with increasing masculinity, brows move downwards in 2D images, but move downwards and protrude in 3D images. Moreover, and related to this latter point, information available from 2D and 3D faces might differ. The additional information available from lateral views in 3D faces might change how information from frontal views is integrated. Systematic differences between 2D and 3D stimuli are further investigated in Appendix E.

Second, differences between preferences in 2D and 3D images found in the current study might be related to a methodological artefact. While the base stimulus faces in the 2D and 3D tasks were matched on age, BMI and height, the 2D and 3D sex prototypes on which masculinity transformations were based were not matched on these variables and might have differed in their dimorphism. That is, it could be that 2D prototypes were more sexually dimorphic than the 3D prototypes. This would mean that moving a 2D face towards 50% higher masculinity would result in a more masculine face shape than applying the same percentage of masculinity to a 3D face. Potential differences in masculinity transforms due to differences in prototypes are further investigated in Appendix F.

Attractiveness as a function of male masculinity Findings from the interactive tasks replicated previous findings of an effect of temporal relationship context on masculinity preferences. Women preferred a higher level of masculinity in a short-term compared to long-term context, which has been previously interpreted as evidence that in a long-term context, high levels of masculinity, associated with less prosocial behaviour, might be more costly than in a short-term context (see Section 9.1.3). An effect of temporal context of masculinity preferences in the rating task was not assessed.

The initial analysis of women's attractiveness ratings as a function of masculinity level showed that, as predicted, men's facial masculinity was related to women's ratings of attractiveness in a curvilinear fashion. Very low and very high levels of masculinity were rated as relatively unattractive. Attractiveness ratings peaked at a level of +87% masculinity.

The preference curves were asymmetric in shape for the presented range of masculinity. Very low levels of masculinity were rated as notably less attractive than very high levels, a pattern not predicted by the extensive literature focussing on the costs associated with (facial) masculinity. One possible explanation for the decreased tolerance of low as compared to high levels of masculinity might be found in the relatively low quality of the texture images with which the stimulus faces were rendered. Base faces might not have been perceived as particularly masculine to begin with. Transforming a face of average masculine shape towards

50% greater femininity renders the face as androgynous in shape; a transformation of −100% effectively corresponds to the face shape of an average woman. Due to the potentially lower starting level in perceived masculinity, the applied transformations towards greater femininity might have been too extreme. Note however, that early work on masculinity preferences used similarly low resolution images (e.g., Perrett et al., 1998), yet found a much lower preferred masculinity level (close to +90% in the current study compared to −10% in Perrett et al., 1998).

Control variables *Age.* In line with previous findings (e.g., Little et al., 2001), I found that older women preferred a higher level of masculinity in the interactive tasks. Facial masculinity has been previously found to increase men's perceived age (Boothroyd et al., 2005; Perrett et al., 1998). Older-, more mature-looking men might be particularly attractive to older women because of the association of age and social status, as well as a tendency in women to prefer somewhat older men than themselves (e.g., Buss, 1989). Consistent with the findings from the interactive task, older women's attractiveness ratings peaked at a slightly higher level than those of younger women in the rating task.

My findings add to the previous observation of an age effect by showing that age not only seems to affect peak preference level, but attractiveness ratings as a function of masculinity more generally—older women showed a higher tolerance for both low and high levels of masculinity. This might reflect a decreased mate value of older women (e.g., Pawlowski & Dunbar, 1999). Younger women gave less favourable ratings on both sides of their peak preference level. At the peak preference level itself, their ratings of attractiveness were actually higher than those of older women's ratings at peak preference level. Taken together, these findings suggest that while younger women, on average, prefer a lower level of masculinity, their ratings of attractiveness are more sensitive to facial masculinity.

Sexual orientation. Analyses from the interactive tasks showed that the more women were exclusively sexually attracted to men, the higher their preferred level of masculinity was. This finding is in line with findings from the preference curves: the more women were exclusively sexually attracted to men, the more attractive they found highly masculine faces. Findings from the preference curves suggest that cost/benefit functions of masculinity do not differ depending on how exclusively women feel attracted to men—rather, the entire preference curve is just shifted towards higher levels of masculinity. Together with findings by Batres et al. (n.d.), this suggests that even within women who might report their sexual orientation as heterosexual if only given the options of homosexual/bisexual/heterosexual, variation in attraction to men affects masculinity preferences, and should therefore be accounted for in future studies.

Hormonal contraceptive use. In contrast to previous findings, my results showed that women using hormonal contraceptives preferred a higher level of masculinity in the interactive tasks than women who did not use hormonal contraceptives. This is surprising given recent findings showing that for women who started using the pill, masculinity preferences shifted toward less masculinity (Little, DeBruine & Jones, 2013). With regards to masculinity preference curves, hormonal contraceptives were found to have no effect.

Base models for subsequent analyses Age was found to affect masculinity preferences in both the interactive as well as rating tasks, and was thus retained as a control variable. Sexual orientation, too, was found to affect preferences in both the interactive and rating tasks. Sample sizes for the different sexual orientation categories were relatively small and unbalanced, making it hard to statistically control for sexual orientation in the analyses of other effects. Thus, it was decided to restrict subsequent analyses to women who identified as completely heterosexual (N=223).[4]

While women using hormonal contraceptives preferred more masculine faces in the interactive task than women not using hormonal contraceptives, hormonal contraceptive use did not affect masculinity preference curves in the rating task. As the aim of this chapter was to identify potentially adaptive tuning of mate preferences and hormonal contraceptive use is an evolutionary novel phenomenon, the following approach was adopted for subsequent analyses: If previous literature suggested that hormonal contraceptive use might interact with the variable of interest in predicting masculinity preferences, women who reported to use hormonal contraceptives were excluded from analyses. If no previous literature existed, an interaction with hormonal contraceptive use was modelled. If an interaction was found, analyses were restricted to women not using hormonal contraceptives.

The base models to which questionnaire items were added as fixed effects were thus the following.

> Interactive task: preferred masculinity level ~ stimulus dimension + temporal context + temporal context x stimulus dimension + age + age x dimension + (1|sample/subject/face).[5]

> Tuning curves: attractiveness ~ masculinity level + masculinity level^2 + age + age x masculinity level + age x masculinity level^2+(1|sample/subject/face)

[4] On the 7-point scale, 23 women indicated their sexual orientation as *5*, 69 as *6* and 223 as *7*, i.e. completely heterosexual.

[5] The symbol "~" should be read as "predicted by". The term "(1|sample/participant/face)" denotes the random effects: a random intercept was modelled for stimulus base face nested in participant nested in sample.

9.3.2. Self-rated attractiveness

9.3.2.1. Results

Four women did not report their self-rated attractiveness, and for one additional women data on hormonal contraceptive use was missing. The sample size was thus N=218. The mean self-rated attractiveness was 4.6±1.2, with values ranging from 1 to 7.

Preferences in interactive tasks The effect of self-rated attractiveness on masculinity preferences was tested by adding self-rated attractiveness to the interactive task base model (see Section 9.3.1.2). To test whether effects might differ for 2D and 3D stimuli, I allowed for an interaction between self-rated attractiveness and stimulus dimension. As effects have been previously shown to be contingent upon the temporal context for which attractiveness is assessed, we also allowed for an interaction between self-rated attractiveness and context. To control for the effects of hormonal contraceptive use, hormonal contraceptive use was added to the model, and three-way interactions between hormonal contraceptive use and self-rated attractiveness x dimension and self-rated attractiveness x context was tested.

The three-way interaction between self-rated attractiveness x dimension x pill use was significant (p=.009). Analyses were thus restricted to women not using hormonal contraceptives (N=156). For women not using hormonal contraceptives, a significant interaction between self-rated attractiveness and relationship context was observed (see Table 9.5). As can be seen from Figure 9.3.9, women who rated themselves as more attractive preferred similar masculinity levels in both contexts, whereas women who rated themselves as less attractive preferred more masculine faces in a short-term compared to a long-term context.

Table 9.5.: Effects of self-rated attractiveness on masculinity preferences in the interactive tasks for women not using hormonal contraceptives. Self-rated attractiveness was reported on a 7-point scale; high values indicate high self-rated attractiveness.

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	4.14	0.35	202	11.85	<.001
context *long-term*	-0.61	0.15	2305	-4.12	<.001
dimension *2D*	-1.33	0.21	2305	-6.31	<.001
dimension *2D* x context *long-term*	0.10	0.08	2306	1.29	.196
age	0.01	0.01	185	1.63	.105
age x dimension *2D*	0.01	0.00	2305	1.61	.108
self_attractive	0.00	0.05	220	0.01	**.992**
self_attractive x context *long-term*	0.10	0.03	2305	3.36	**.001**
self_attractive x dimension *2D*	0.01	0.03	2305	0.18	**.861**

Preferences in rating task: tuning of preferences To control for the effects of hormonal contraceptive use, self-rated attractiveness was entered in a first step with hormonal contraceptive use into the tuning curve base model (see Section 9.3.1.3). The three-way interaction between hormonal contraceptive use, self-rated attractiveness and masculinity level was significant for both the linear (p=.002) and quadratic masculinity terms (p=.004). Analyses were thus restricted to women not using hormonal contraceptives.

The model revealed a significant main effect of self-rated attractiveness on ratings of male attractiveness, as well as significant interactions with both the linear and quadratic masculinity level terms (see Table 9.6). Figure 9.3.10 visualizes preference curves for women who rated themselves lowest (1) and highest (7) on attractiveness.

Table 9.6.: Effect of self-rated attractiveness on masculinity preference curves in women not using hormonal contraceptives. Self-rated attractiveness was reported on a 7-point scale; high values indicate high self-rated attractiveness.

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	2.73	0.63	48	4.32	<.001
mlevel	0.62	0.23	3738	2.64	.008
mlevelsq	-0.07	0.03	3738	-2.50	.013
age	0.01	0.01	367	0.61	.544
age x mlevel	0.00	0.01	3738	-0.78	.436
age x mlevelsq	0.00	0.00	3738	0.74	.460
self_attractive	-0.24	0.09	513	-2.80	**.005**
self_attractive x mlevel	0.11	0.03	3738	3.05	**.002**
self_attractive x mlevelsq	-0.01	0.00	3738	-2.34	**.019**

9.3.2.2. Discussion

In line with previous findings (Little et al., 2001; Penton-Voak et al., 2003), results from the interactive tasks revealed an interaction between women's self-rated attractiveness and the type of relationship for which men's faces were assessed. Women who perceived themselves as less attractive shifted their masculinity preferences towards lower levels of masculinity in a long-term compared to a short-term relationship context. In contrast, women who perceived themselves as more attractive preferred a similar level of masculinity in both short- and long-term contexts. This has been previously taken as evidence that less attractive women have a lesser mate value and thus might opt for long-term mates who are "more likely to invest or least likely to desert" (Little et al., 2001, p. 42,)—i.e. it has been suggested that the

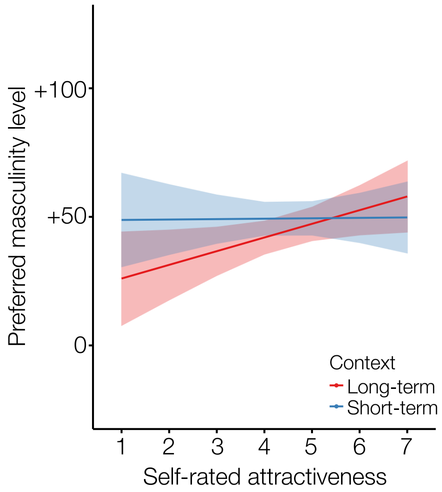


Figure 9.3.9.: Women rating themselves as less attractive preferred less masculine faces in a long-term compared to a short-term relationship context, while women rating themselves as more attractive showed similar preferences in both contexts. Bands represent the respective 95% CI.

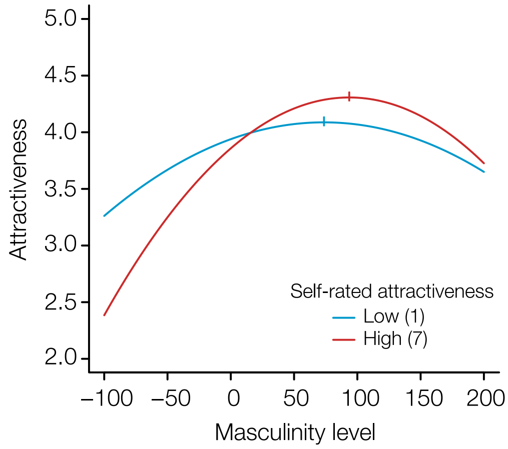


Figure 9.3.10.: Compared to women who rated themselves as less attractive, women who rated themselves as more attractive showed a higher masculinity peak preference level (94% vs 74%), and less tolerance towards lower levels of masculinity.

cost of choosing a highly masculine partner might be higher for women with low mate value compared to women with a higher mate value.

Results from the rating task suggest differently—preference curves show that both women of self-rated low and high attractiveness show a similar level of attraction to highly masculine men. It seems that it is not the costs of *high* levels of masculinity that differ for women depending on their perceived individual condition; instead, there seems to be a difference in the costs associated with choosing a very feminine mate. More attractive women showed less tolerance to *low* levels of masculinity than less attractive women. Attractiveness ratings of more attractive women were lower for the lowest level of masculinity, and increased more steeply with increasing masculinity. More generally, more attractive women showed a greater variance in their ratings of attractiveness, which might indicate they are more discriminatory when it comes to men's facial masculinity.

9.3.3. Relationship status

9.3.3.1. Results

Based on findings from Little et al. (2002), analyses were restricted to women not using hormonal contraceptives (N=157). Relationship status was missing for 11 women; the final sample size was thus N=146 (59 single women, 87 partnered women).

Preferences in interactive tasks The effect of relationship status on masculinity preferences was tested by adding relationship status to the interactive task base model (see Section 9.3.1.3). To test whether effects might differ for 2D and 3D stimuli, I allowed for an interaction between relationship status and stimulus dimension. As previous studies suggested that the effect of relationship status might be contingent on the type of relationship for which attractiveness is assessed (Little et al., 2002), we also tested for an interaction between relationship status and temporal context.

Relationship status had no main effect on masculinity preferences, but interacted with stimulus dimension (see Table 9.7). As Figure 9.3.11 shows, when presented with 2D stimuli, partnered women preferred a higher level of masculinity than single women, while this difference was less pronounced for 3D faces. The interaction between relationship status and relationship context was not significant.

Table 9.7.: Effect of relationship status on masculinity preferences in interactive tasks in women not using hormonal contraceptives. The reference category for relationship status is "partner *no*".

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	4.13	0.27	17	15.53	<.001
context *long-term*	-0.18	0.07	2156	-2.40	.017
dimension *2D*	-1.24	0.16	2155	-7.79	<.001
dimension *2D* x context *long-term*	0.09	0.08	2156	1.10	.273
age	0.01	0.01	51	1.39	.170
age x dim *2D*	0.00	0.01	2155	0.04	.968
partner *yes*	0.00	0.16	186	0.03	**.977**
partner *yes* x dim *2D*	0.26	0.09	2156	2.88	**.004**
partner *yes* x context *long-term*	0.09	0.08	2156	1.13	**.261**

Preferences in rating task: tuning of preferences Adding relationship status to the tuning curve base model (see Section 9.3.1.2) revealed that relationship status had neither a main effect on ratings of male attractiveness nor did it interact with masculinity level terms in predicting women's ratings of attractiveness (see Table 9.8).

9.3.3.2. Discussion

In line with previous findings, single women were found to prefer less masculine men than partnered women when tested with 2D stimuli. This finding has been previously interpreted as

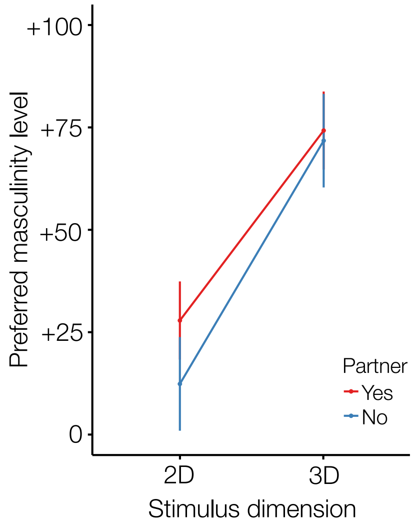


Figure 9.3.11.: Partnered women preferred a higher level of masculinity in the interactive 2D task but less so in the interactive 3D task.

Table 9.8.: Effect of relationship status on masculinity preference curves in women not using hormonal contraceptives. The reference category for relationship status is "partner *no*".

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	1.68	0.52	13	3.25	.006
mlevel	1.07	0.18	3498	6.05	<.001
mlevelsq	-0.11	0.02	3498	-5.17	<.001
age	0.01	0.02	434	0.62	.538
age x mlevel	0.00	0.01	3498	-0.04	.967
age x mlevelsq	0.00	0.00	3498	-0.12	.902
partner *yes*	-0.23	0.25	489	-0.90	**.371**
partner *yes* x mlevel	-0.11	0.10	3498	-1.07	**.286**
partner *yes* x mlevelsq	0.02	0.01	3498	1.36	**.175**

evidence that compared to single women, for partnered women mating with a more masculine men has higher benefits (cf. immunocompetence handicap hypothesis) and fewer costs (less prosocial behaviour of masculine men being less relevant in short-term context). However, no evidence for such a tuning of preferences was observed in the 3D rating task, and the relationship status effect was less pronounced in the interactive task when 3D faces were assessed.

9.3.4. Financial worries

9.3.4.1. Results

Based on findings from Little, Cohen et al. (2007), analyses were restricted to women not using hormonal contraceptives (N=157). One woman did not report her financial worries; the final sample size was thus N=*156*. The mean reported financial worries was 4.5±1.8, with values ranging from 1 to 7.

Preferences in interactive tasks The effect of financial worries on masculinity preferences was tested by adding financial worries to the interactive task base model (see Section 9.3.1.3). To test whether effects might differ for 2D and 3D stimuli, I allowed for an interaction between financial worries and stimulus dimension. As effects of environmental harshness have been previously shown to be contingent upon the temporal context for which attractiveness is assessed, I also allowed for an interaction between financial worries and context.

The tested model revealed no significant main effect of, or interactions with, financial worries (see Table 9.9).

Table 9.9.: Effects of financial worries on masculinity preferences in interactive tasks in women not using hormonal contraceptives. Financial worries were recorded on a 7-point scale; high values indicate many financial worries.

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	4.20	0.29	198	14.62	<.001
context *long-term*	-0.26	0.11	2306	-2.36	.019
dimension *2D*	-1.32	0.18	2305	-7.52	<.001
dimension *2D* x context *long-term*	0.08	0.08	2306	1.06	.291
age	0.01	0.01	184	1.49	.138
age x dimension *2D*	0.01	0.00	2305	1.59	.112
worries	-0.01	0.04	220	-0.26	**.793**
worries x dimension *2D*	0.01	0.02	2306	0.30	**.765**
worries x context *long-term*	0.03	0.02	2306	1.22	**.224**

Preferences in rating task: tuning of preferences Adding financial worries to the tuning curves base model showed no main effect of financial worries on attractiveness ratings, but a significant interaction between financial worries and the quadratic masculinity term (see Table 9.10). Figure 9.3.12 visualizes masculinity preference curves for women who reported the least (1) and most (7) financial worries.

Table 9.10.: Effect of financial worries on masculinity preference curves in women not using hormonal contraceptives. Financial worries were reported on a 7-point scale; high values indicate many financial worries.

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	1.71	0.55	23	3.12	.005
mlevel	1.02	0.19	3738	5.31	<.001
mlevelsq	-0.09	0.02	3738	-3.97	<.001
age	0.01	0.01	388	0.75	.453
age x mlevel	-0.01	0.01	3738	-1.05	.296
age x mlevelsq	0.00	0.00	3738	0.97	.334
worries	-0.03	0.06	515	-0.57	**.570**
worries x mlevel	0.03	0.02	3738	1.16	**.247**
worries x mlevelsq	-0.01	0.00	3738	-2.15	**.032**

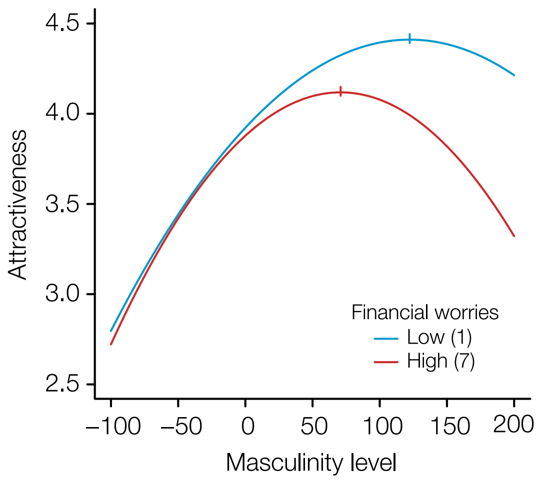


Figure 9.3.12.: Women who reported fewer financial worries showed a higher peak preference than women who reported more financial worries (122% vs 71%) and a higher tolerance for high levels of masculinity. The attractiveness of low levels of masculinity appeared unaffected by individual differences in financial worries.

9.3.4.2. Discussion

We asked participants how much they worried about their future financial situation to approximate women's perceived environmental harshness. While we found no effect on masculinity preferences in the interactive task, findings from the rating task are in line with previous research on the effect of perceived environmental harshness: women in perceived harsh environments preferred less masculine men than women in safe environments (Little, Apicella & Marlowe, 2007; Little, DeBruine & Jones, 2013), reflected in a lower peak preference level. Our findings add to previous observations by showing that women with less financial worries preferred very masculine men; respectively, women with more financial worries appeared to dislike very masculine men rather than preferring low levels of masculinity—very feminine faces were rated as equally (un)attractive by women with little and high financial worries.

Moore et al. (2006) found that women in control of financial resources placed greater importance on men's physical attractiveness—from the current work, this may translate to greater attraction to high levels of masculinity. Note that this interpretation is in conflict with previous studies which suggested environmental harshness would be linked to a preference for facial cues to male pro-sociality, i.e. less masculine men.

9.3.5. Pathogen disgust sensitivity

9.3.5.1. Results

Ten participants did not respond to all pathogen disgust questionnaire items. The final sample size was thus N=213. The mean pathogen disgust sensitivity was 25.4±7.7, with values ranging from 2 to 42.

Preferences in interactive tasks In a first step, I tested for an interaction between hormonal contraceptive use and pathogen disgust sensitivity, as well as disgust sensitivity and age (Lee & Zietsch, 2015) by adding the following terms to the interactive task base model (see Section 9.3.1.3): *disgust + hormonal contraceptive use + disgust x stimulus dimension + disgust x age + disgust x hormonal contraceptive use*

Neither the main effect of contraceptive use, nor the interaction between contraceptive use and pathogen disgust sensitivity were significant (both p>.436). Omitting hormonal contraceptive use from the model revealed a trend for an interaction between age and disgust (see Table 9.11). While younger women's masculinity preferences increased with increasing pathogen disgust sensitivity, older women's masculinity preferences decreased with increasing pathogen disgust sensitivity (see Figure 9.3.13).

Table 9.11.: Effects of pathogen disgust sensitivity on masculinity preferences in interactive tasks. Pathogen disgust sensitivity ranged from 2 to 42, with high values indicating high pathogen disgust sensitivity.

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	3.00	0.79	214	3.78	<.001
context *long-term*	-0.15	0.05	3153	-3.19	.001
dimension *2D*	-1.08	0.17	3154	-6.22	<.001
dimension *2D* x context *long-term*	0.09	0.07	3154	1.33	.184
age	0.06	0.02	211	2.30	.022
age x dimension *2D*	0.01	0.00	3153	1.38	.169
disgust	0.05	0.03	211	1.55	**.124**
disgust x dimension *2D*	0.00	0.00	3156	-1.07	**.284**
disgust x age	0.00	0.00	208	-1.91	**.057**

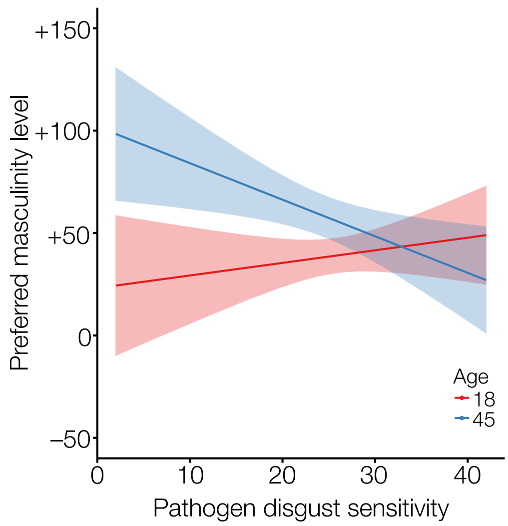


Figure 9.3.13.: A trend for an interaction between disgust and age was found. For younger women, increasing pathogen disgust sensitivity was associated with increasing masculinity preferences, while for older women the opposite was observed. Bands represent the respective 95% CI.

Preferences in rating task: tuning of preferences Following findings by Lee and Zietsch (2015) and findings from the interactive tasks, I conducted analyses separately for younger and older women by splitting the sample by median age (30 years).

In younger women, the interaction between pathogen disgust sensitivity and the linear masculinity term was significant, while the main effect of pathogen disgust sensitivity and the interaction with the quadratic masculinity term were not significant (see Table 9.12). Figure 9.3.14 shows the preference curves for minimum and maximum pathogen disgust sensitivity at the subsample median age (23 years).

Table 9.12.: Effects of pathogen disgust sensitivity on masculinity preferences curves. Pathogen disgust sensitivity ranged from 2 to 42, with high values indicating high pathogen disgust sensitivity.

Sample		estimate	*SE*	df	*t*-value	*p*
30 or younger	(Intercept)	0.50	1.04	75	0.48	.633
(*N*=111)	mlevel	1.46	0.40	2658	3.63	<.001
	mlevelsq	−0.19	0.05	2658	−3.82	<.001
	age	0.10	0.03	288	2.75	.006
	age x mlevel	−0.04	0.01	2658	−2.82	.005
	age x mlevelsq	0.01	0.00	2658	3.17	.002
	disgust	−0.03	0.02	507	−1.81	**.071**
	disgust x mlevel	0.02	0.01	2658	2.26	**.024**
	disgust x mlevelsq	−0.00	0.00	2658	−1.66	**.097**
31 and older	(Intercept)	1.28	1.24	294	1.04	.302
(*N*=102)	mlevel	0.98	0.48	2442	2.05	.041
	mlevelsq	−0.04	0.06	2442	−0.75	.456
	age	0.04	0.03	294	1.18	.237
	age x mlevel	−0.01	0.01	2442	−0.59	.559
	age x mlevelsq	0.00	0.00	2442	−0.34	.733
	disgust	−0.03	0.01	294	−1.86	**.064**
	disgust x mlevel	0.01	0.01	2442	1.68	**.094**
	disgust x mlevelsq	−0.00	0.00	2442	−2.03	**.042**

In older women, the interaction between pathogen disgust sensitivity and the quadratic masculinity term was significant, while the main effect and the interaction with the linear term were not significant (see Table 9.12). Figure 9.3.15 shows the preference curves for minimum and maximum pathogen disgust sensitivity at the subsample median age (37 years).

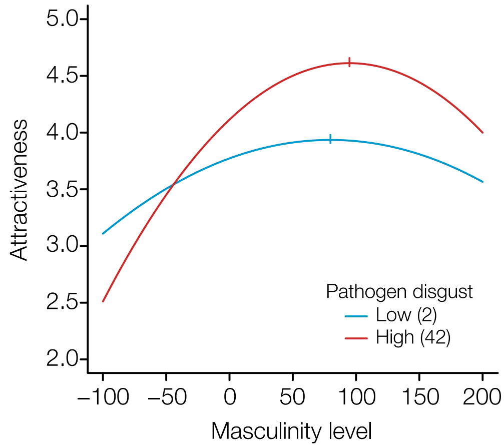


Figure 9.3.14.: In the younger subsample, women with high pathogen disgust sensitivity rated the least masculine faces as less favourable than women with low pathogen disgust sensitivity, and showed a higher increase in ratings of attractiveness with increasing masculinity.

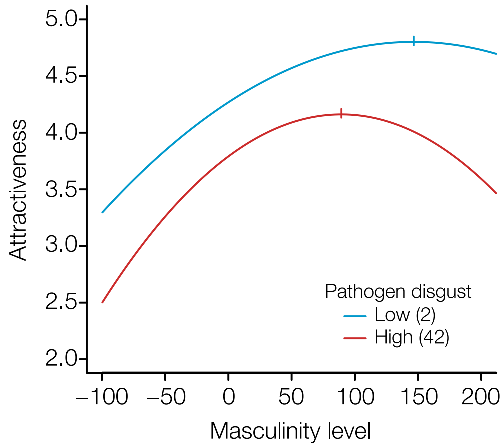


Figure 9.3.15.: In the older subsample, women with low pathogen disgust showed a higher peak preference level than women with high pathogen disgust sensitivity and their preferences remained high beyond the peak level. For women with high pathogen disgust sensitivity, attractiveness ratings decreased more strongly after reaching peak preference.

9.3.5.2. Discussion

In contrast to Lee and Zietsch (2015), I found a marginally significant interaction between participant age and pathogen disgust sensitivity in predicting masculinity preferences in the interactive task: while younger women's masculinity preferences increased with increasing pathogen disgust sensitivity (in line with, e.g., DeBruine, Jones, Tybur et al., 2010), older women's masculinity preferences decreased with increasing pathogen disgust sensitivity. I note that while Lee and Zietsch (2015, Study 1) tested preferences in a similar sample to the current one (heterosexual MTurk users residing in the US), their participants were considerably—i.e. on average seven years—older than mine. The effect they reported, while not significant, was comparable to the one I observed (Lee and Zietsch: t=−1.40, current study t=−1.91). This (trend for an) interaction between age and pathogen disgust can account for the null finding in Lee and Zietsch' overall sample, as well as their finding of a positive relation between pathogen disgust sensitivity and masculinity preferences in the younger subsample. It supports their interpretation that the effect of pathogen disgust on masculinity preferences is age-dependent. Younger women who are worried about pathogen contamination appear to prefer men with more masculine faces, and this may be due to the heritable good health associated with masculinity (DeBruine, Jones, Tybur et al., 2010). It remains to be clarified why the association of pathogen disgust sensitivity and masculinity preference in the interactive task was found to be reversed in older age.

In line with findings from the interactive task, I found that attractiveness ratings of younger women who score higher on pathogen disgust sensitivity show a greater sensitivity to facial masculinity than those of women who score lower on pathogen disgust sensitivity. While for women with low pathogen disgust sensitivity the preference curve was relatively flat, women with high pathogen disgust sensitivity showed a greater variation in their attractiveness ratings. They rated low levels of masculinity as less attractive than women with low pathogen disgust sensitivity, but their attractiveness ratings increased more steeply with increasing masculinity, peaked at a higher level of masculinity, and their attractiveness ratings at that peak level were higher than attractiveness ratings at the peak level of women with low pathogen disgust sensitivity.

In older women, a different pattern was observed. Here, both women with low and high pathogen disgust sensitivity showed a relatively steep increase in attractiveness ratings with increasing masculinity. Older women with high pathogen disgust sensitivity showed a similar preference curve to younger women with high pathogen disgust sensitivity. However, in contrast to younger women, older women with low pathogen disgust sensitivity gave more favourable attractiveness ratings in general, and showed a higher peak preference level than women with high pathogen disgust sensitivity; indeed, their preference curve attenuated minimally after the peak preference level.

9.3.6. Self-reported health

9.3.6.1. Results

As self-reported health might be related to self-perceived condition, which has been found to interact with pill use in predicting masculinity preferences, analyses were restricted to women not using hormonal contraceptives (N=157). One woman did not report her health; the final sample size was thus N=156. The mean self-reported health was 2.3±0.9, with values ranging from 1 to 4.

Preferences in interactive tasks Self-reported health was added to the interactive task base model and allowed to interact with both context and stimulus dimension. The main effect of self-reported health was not significant; neither was the interaction with stimulus dimension (see Table 9.13). Self-reported health did interact significantly with temporal relationship context. As can be seen from Figure 9.3.16, women who reported better health did not show differences in masculinity preferences for short- and long-term contexts, whereas women who reported worse health showed a lower preference for masculinity in a long-term compared to a short-term context.

Table 9.13.: Effects of self-rated health on masculinity preferences in the interactive tasks for women not using hormonal contraceptives. Self-rated health was reported on a 5-point scale; high values indicate good self-rated health.

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	4.21	0.38	203	11.03	<.001
context *long-term*	-0.46	0.17	2305	-2.71	.007
dimension *2D*	-1.30	0.23	2305	-5.69	<.001
dimension *2D* x context *long-term*	0.09	0.08	2306	1.21	.225
age	0.01	0.01	184	1.58	.116
age x dimension *2D*	0.01	0.00	2305	1.61	.108
health	-0.02	0.07	219	-0.27	**.789**
health x dimension *2D*	0.00	0.04	2306	0.05	**.960**
health x context *long-term*	0.08	0.04	2305	1.99	**.047**

Preferences in rating task: tuning of preferences Adding self-reported health to the base model showed a significant negative main effect of self-reported health on attractiveness ratings (see Table 9.14). Neither of the interactions with self-reported health was significant Figure 9.3.17 visualizes masculinity preference curves for women who reported to be of worst (1) and best (5) health.

Table 9.14.: Effect of self-rated health on masculinity preference curves in women not using hormonal contraceptives. Self-rated health was reported on a 5-point scale; high values indicate good self-rated health.

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	2.64	0.67	65	3.95	<.001
mlevel	0.94	0.25	3738	3.71	<.001
mlevelsq	-0.13	0.03	3738	-4.06	<.001
age	0.01	0.01	358	0.63	.528
age x mlevel	0.00	0.01	3738	-0.89	.373
age x mlevelsq	0.00	0.00	3738	0.92	.360
health	-0.27	0.12	499	-2.26	**.024**
health x mlevel	0.05	0.05	3738	1.02	**.306**
health x mlevelsq	0.00	0.01	3738	0.23	**.822**

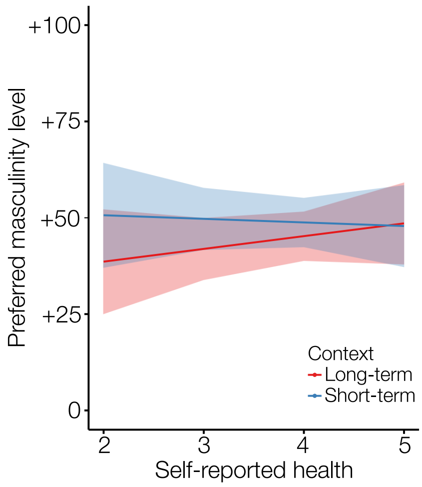


Figure 9.3.16.: Self-reported health interacted with relationship context in predicting masculinity preferences: While women who reported poorer health showed a higher preference for masculinity in a short-term compared to long-term relationship context, women who reported better health showed similar masculinity preferences in both contexts. Bands represent the respective 95% CI.

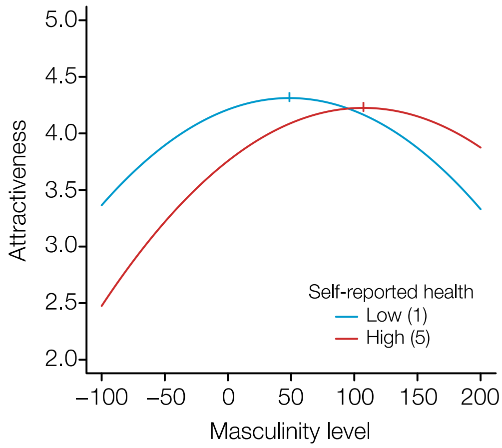


Figure 9.3.17.: Women who reported better health showed a higher peak preference level and less tolerance towards lower compared to higher level of masculinity. Women who reported poorer health showed a similar tolerance towards low and high levels of masculinity.

9.3.6.2. Discussion

My findings are in conflict with Feinberg et al. (2012) and De Barra et al. (2013), who concluded that (self-reported) health was negatively correlated with masculinity preferences in faces (and voices), and that effects of self-reported attractiveness and health have dissociable effects on masculinity preferences. The current study found that both self-rated health and attractiveness were positively related to masculinity preferences in a long-term relationship context.

My findings point to self-reported health affecting condition/mate value rather than affecting masculinity preferences due to health benefits associated with choosing a masculine partner. In line with Scott et al. (2008), I found that preferences for masculinity did not differ for women of better and poorer health in a short-term relationship context, but were lower for women with poorer compared to better self-reported health in a long-term relationship context. In support of an association of self-reported health and self-perceived mate value, I found that self-reported health and self-rated attractiveness were moderately correlated (Pearson's $r(271)=.41$, 95% CI [.31 .51], $p<.001$).

Masculinity preference curves suggest that women shift their preferences towards higher levels of masculinity with increasing health; the peak masculinity preference level was higher for women of better compared to poorer health. For women of poorer self-reported health, attractiveness ratings were similar for the lowest and highest presented levels of masculinity, while women of better health appeared to show less tolerance towards lower compared to higher levels of masculinity.

9.3.7. Combining different factors into one model

Previous literature has mainly focussed on investigating how single factors affect masculinity preferences. In a last step, I tested whether established effects on masculinity preference curves would remain significant if combined into one model, as the different established parameters affecting masculinity preferences could reflect a single underlying cause or construct. Analysis was restricted to women who reported not to use hormonal contraceptives.

Results The following terms were added to the base model: relationship status, self-rated attractiveness, financial worries, pathogen disgust sensitivity and self-reported health. All terms were allowed to interact with both the linear and quadratic masculinity terms.

Table 9.15 summarises the results. The effects for self-rated attractiveness, financial worries and pathogen disgust remained significant; effects of relationship status was again not significant, and the effect of self-reported health no longer reached significance.

Table 9.15.: Simultaneously testing for the effect of multiple individual differences on masculinity preference curves in women not using hormonal contraceptives.

	estimate	*SE*	df	*t*-value	*p*
(Intercept)	4.67	0.98	92	4.78	<.00
mlevel	-0.11	0.37	3298	-0.29	.776
mlevelsq	0.01	0.05	3298	0.24	.813
age	0.00	0.02	383	0.15	.882
age x mlevel	0.00	0.01	3298	0.13	.894
age x mlevelsq	0.00	0.00	3298	-0.03	.974
partner	-0.19	0.26	428	-0.73	.467
partner x mlevel	-0.09	0.11	3298	-0.83	.407
partner x mlevelsq	0.01	0.01	3298	0.93	.351
self_attractive	-0.18	0.11	440	-1.61	.109
self_attractive x mlevel	0.10	0.04	3298	2.24	**.025**
self_attractive x mlevelsq	-0.01	0.01	3298	-2.63	**.009**
worries	-0.09	0.07	440	-1.22	.222
worries x mlevel	0.06	0.03	3298	2.05	**.040**
worries x mlevelsq	-0.01	0.00	3298	-2.51	**.012**
disgust	-0.03	0.01	438	-1.76	.079
disgust x mlevel	0.01	0.01	3298	1.87	.062
disgust x mlevelsq	0.00	0.00	3298	-2.09	**.036**
health	-0.25	0.16	433	-1.56	.120
health x mlevel	0.04	0.06	3298	0.62	.535
health x mlevelsq	0.01	0.01	3298	0.66	.511

Discussion My findings suggest that the effects of self-rated attractiveness, perceived environmental harshness and pathogen disgust reported here and elsewhere make independent contributions to masculinity preferences. Self-reported health and attractiveness, however, may be two facets of condition or mate value.

9.4. Tuning curves—worth the effort?

"Point measures" of masculinity preferences, such as interactive or two-alternative forced choice tasks can reveal how individual differences shift preferences towards lower or higher masculinity. The study presented in this chapter was the first to test how women's masculinity preferences for men's faces change across *multiple* levels of masculinity.

First, I established that—as previously assumed but not explicitly tested—male facial attractiveness is related to facial masculinity in a curvilinear fashion. Second, I investigated whether individual differences change the shape of masculinity preference curves. Put differently, we tested whether women differentially tune their masculinity preferences. My approach has shown that individual differences do indeed change the tolerance towards low and high levels of masculinity as previously suggested. For some parameters, preferences do not simply shift towards lower/higher masculinity levels.

As the current study introduced several new methodological aspects compared to previous studies (3D stimuli, rating of different masculinity levels), I first tested whether previously found effects replicated in the current data set when employing previously used methods.

Using 2D and 3D stimuli in an interactive task, I replicated findings for a number of variables: partnered women preferred higher levels of masculinity than single women; less attractive women shifted their preferences towards lower masculinity levels in a long-term context while more attractive women did not; a (perceived) harsh financial environment led to preferences for lower levels of masculinity; younger women's pathogen disgust sensitivity was positively associated with masculinity preferences. I also tested for effects of self-reported health on masculinity preferences and found similar results as for self-reported attractiveness.

Modelling and plotting tuning curves allowed for a more refined investigation of these effects. I was able to show that some individual differences affect the tolerance towards high but not low levels of masculinity (perceived environmental harshness), while others affect the tolerance towards low but not high levels of masculinity (self-rated attractiveness). Other factors were found to shift masculinity preference curves as a whole without changing their shape (exclusive sexual attraction to men, self-reported health). Yet other findings showed changes in sensitivity to masculinity with tuning curves becoming sharper (age, pathogen disgust sensitivity).

These findings are important because of their theoretical implications: for example, they suggest that self-perceived condition does not change the benefits associated with choosing a very masculine mate; rather, self-rated attractiveness appears to change the costs of choosing a very feminine mate. For less attractive women it seems less costly to choose a very feminine men than for more attractive women. This might make it necessary to reconsider previous narratives which have argued that cues to pro-sociality (less masculinity/higher femininity) are actively preferred by less attractive women. I note that the observed comparatively small difference in peak preference levels in the current study is not necessarily in conflict with previous findings: the steeper tuning function for low levels of masculinity in more attractive women can lead to the finding of a large preference difference when using a two-alternative forced choice task using stimuli that range between ±50%, i.e. a low range of masculinity. The use of preference curves instead of point measure also showed that older women and less attractive women seem less discriminatory when it comes to men's facial masculinity—a new finding that should be taken into account when investigating masculinity preferences.

Investigating tuning curves requires participants to rate a set of faces with limited variation in appearance. One disadvantage thus lies in the repetitiveness of the task—participant motivation might be lower than in, for example, an interactive task. Ideally, I would have presented more base faces, at more masculinity levels, under both a short- and long-term relationship context. Due to concerns about attrition rates (especially for the unpaid *online* sample), I decided to keep the task relatively brief. Despite these concerns, I suggest that the use of tuning curves is indeed beneficial.

Part IV.

Facial Cues to Strength

10. Section outline[1]

Recent literature has suggested that facial impressions of dominance and strength have a profound effect on interpersonal perception. Yet, little is known on how physical strength is reflected in facial morphology, and which facial cues observers use to form impressions of strength. This section presents four empirical studies. Study 1 tests how accurately strength is perceived from colour- and texture-standardised 3D faces. Studies 2 and 3 test the hypothesis that perceptions of strength are informed by facial cues to physical characteristics, specifically BMI, muscle and fat mass. As Study 3 introduces new face-morphological scores of body muscularity and fat mass, Study 4 tests whether the shape dimensions described by these scores are perceptually distinct.

[1]Part of the work in this section is currently under review: Holzleitner, I. J., & Perrett, D. I. (under review). Perception of strength from 3D faces is linked to facial cues to physique. *Evolution and Human Behavior*.

11. Introduction

A growing body of literature suggests that intrasexual selection pressures amongst men might have played a more important role in shaping men's traits than has been hitherto acknowledged (Puts, 2010; Puts et al., 2012; Scott et al., 2010). Intrasexual competitiveness, i.e. the drive to compete with other men and the ability to do so successfully, is linked to higher social status, which in turn has positive fitness payoffs (von Rueden, Gurven & Kaplan, 2011). Both intrasexual competitiveness and social status have been argued to be partly based on strength, and in particular upper-body strength, which is tightly linked to fighting ability (Sell et al., 2009). Handgrip strength is a good predictor of upper-body strength (Sell et al., 2009) and overall muscle strength (Wind et al., 2010), and has been found to be associated with behavioral tendencies (such as a propensity for anger and aggressive behaviour Gallup, White & Gallup Jr, 2007; Munoz-Reyes, Gil-Burmann, Fink & Turiegano, 2012; Sell et al., 2009) as well as to influence interpersonal perception (such as impressions of dominance, e.g., Fink et al., 2007).

Sell et al. (2009) emphasized the importance of being able to assess potential rivals' formidability accurately in order to avoid costs from physical conflicts that cannot be won. Similarly, Puts (2010) and Puts et al. (2012) suggested that men's face shape may have developed to signal the ability to successfully engage in competitive encounters to potential rivals. Although it could also be argued that observers learn any consistent cues to strength, the impact of facial impressions of dominance and strength on interpersonal perception indeed seems to be profound. Oosterhof and Todorov (2008), for example, have argued that faces are assessed on two main dimensions, one of which is based on facial cues to physical strength (i.e. the dominance or power dimension, revealing the ability to inflict damage on others as opposed to the valence dimension, which reveals pro- or antisocial intentions). In line with the proposed importance of visual cues to strength in social interactions, Sell et al. (2009) showed that observers can judge men's upper-body strength accurately from facial images alone. They did not, however, investigate which facial cues underpin such judgments.

Recent papers have investigated how strength is reflected in face shape, and which facial features might be driving judgments of strength and formidability. By regressing handgrip strength on two-dimensional (2D) face shape, Windhager et al. (2011) found that strength is associated with a rounder facial shape, a widening between eyebrows, a shorter nose, broadening of the lower face and pronounced jaw muscles (masseter region). Toscano, Schubert and Sell (2014) tested which facial features—used by Zebrowitz et al. (2003) and Zebrowitz, Kikuchi and Fellous (2007)—were associated with the perception of strength and found that faces with a lower eyebrow height, a shorter eye length (i.e. less opened/smaller eyes) and a wider nose were perceived as both stronger and more dominant. Yet, it remains unclear why these features may be related to perceptions of strength and dominance. Recently, Zilioli et al. (2014) identified a facial cue that may mediate perceptions of formidability: facial width to height ratio (fWHR) was linked to both actual fighting ability as well as perceived formidability. fWHR may be linked to formidability through an association with physical strength, or through its association with a propensity for aggressive behavior (e.g., Carré & McCormick, 2008; Carré, McCormick & Mondloch, 2009), although these explanations are not necessarily mutually exclusive given the link between strength and aggression.

Here, we aimed to test whether perceptions of strength might be mediated by facial cues to body physique. That is, instead of pre-defined face features, we investigated whether global variation in face shape linked to body parameters can explain perceptions of strength from faces. If it is adaptive to perceive strength accurately in order to assess fighting ability (Sell et al., 2009), judgments of strength should be based on facial cues to physical characteristics that predict actual strength. Thus, we investigated whether anthropometric variables that relate to actual strength are reflected in face shape, and hypothesized that face shape associated with physical predictors of actual strength would contribute to the perception of strength.

Four studies were conducted. Study 3.1 tested whether strength could be perceived accurately from color- and texture-standardized 3D faces, and visualized the facial correlates of actual and perceived strength. Studies 3.2 and 3.3 investigated which physical parameters are predictive of strength, how they are reflected in face shape, and whether facial correlates of body physique predict perceived strength.

Most previous studies have investigated anthropometric predictors of strength within a clinical context. Two of the most basic descriptors of body physique that are positively correlated with (handgrip) strength are body mass index (BMI, weight[kg]/height[m^2]) and height (e.g., Balogun, Akinloye & Adenlola, 1991; Chandrasekaran, Ghosh, Prasad, Krishnan & Chandrasharma, 2010; Fink, Weege, Manning & Trivers, 2014; Sartorio, Lafortuna, Pogliaghi & Trecate, 2002). Section 8.3 showed that facial cues to BMI and height can be relatively simply assessed and used in a model to explain perceptual ratings of masculinity. In Study 3.2, we thus tested whether facial cues to BMI and height are predictors of perceptual ratings of strength.

While BMI is associated with strength, it conflates muscle mass and fat mass. Perhaps counterintuitively, muscle and fat mass are positively correlated. A weight gain due to nutritional intake leads to an increase in both body fat and lean body mass, potentially due to muscle hypertrophy as a result of increased weight bearing (Forbes, 1987, 1993). This increase in lean mass with weight gain appears to be to some extent sex-specific: at least in obese samples, lean mass increased more strongly with increasing weight in men and boys compared to women and girls (Lafortuna, Maffiuletti, Agosti & Sartorio, 2005; Sartorio, Agosti, De Col & Lafortuna, 2006; Sartorio et al., 2004). In essence, being heavier results in higher *absolute* strength (Sartorio et al., 2006, 2004), reflected in findings that obese participants have higher (anaerobic) strength than a normal-weight control group (Lafortuna et al., 2005), and reflected by the general positive association of weight/BMI and strength (compare weight classes in sporting events).

Despite the correlation of lean and fat mass, underlying body composition in terms of fat and muscle may be a better predictor of strength than BMI for two reasons. First, at a given BMI level, the amount of lean mass can differ. For example, Deurenberg, Yap and van Staveren (1998) reported that, at the same BMI level, European Caucasians have a higher percentage body fat than American Caucasians. Moreover, while fat and muscle appear to be positively correlated when it comes to nutrition-related weight gains, androgens such as testosterone are associated with an increase in lean body mass, but a decrease in fat mass (e.g., Bhasin, Woodhouse & Storer, 2003; Forbes, 1993). Hence, despite having the same BMI, men can differ in their muscle mass and thus in their strength. Second, while being heavier will usually result in being stronger in absolute terms, body fat has a negative impact on muscle quality or *relative* strength, i.e. strength scaled to body or muscle mass (Goodpaster et al., 2001; Newman et al., 2003; Vilaca et al., 2014; Zhang, Peterson, Su & Wang, 2015). Indeed, Sartorio et al. (2002) found that controlling for BMI, lean mass is the best predictor of grip strength in a sample of healthy children, while percentage body fat was negatively related to grip strength. Thus, if two men have the same BMI, but differ in their proportion of lean to fat mass, the man with the higher proportion of muscle mass will be stronger; or, put differently, at the same level of lean mass, having more body fat will negatively affect strength.

In Study 3.3, we tested whether facial cues to muscle and fat could be separated and whether they relate to the perception of strength. We predicted that facial cues to muscle mass would positively predict perceptions of strength. As the relationship of fat mass and strength is complex (positive association with absolute strength, negative relation to relative strength), we predicted facial cues to body fat would impact on perceived strength, but made no prediction regarding the direction of this effect.

In summary, Studies 3.2 and 3.3 had the following research questions.

1. Do anthropometric variables (BMI/height, muscle/fat mass) predict strength?

2. Do these anthropometric parameters relate to face shape?

3. Do facial estimates of anthropometric parameters predict perceptions of strength?

To our knowledge, muscle and fat mass have not been separately related to 3D face shape before. Study 3.4 thus tested whether the face shape associated with fat and muscle would be perceived as being related to body fat and body muscularity, and whether these two dimensions would be perceptually distinguishable from each other.

12. General material and methods

12.1. Stimulus dataset

12.1.1. 3D Images

The facial images of Face Set 2 (see Section 4.2) served as stimulus images.

12.1.2. Anthropometric measurements

After removing footwear and excess clothing, participants' height was measured and weight and body composition (muscle and fat mass) were assessed barefoot using an electrical impedance scale (Tanita SC–330). Height and weight were recorded for all participants, but body composition measures could not be accurately assessed due to the wearing of tights for 10 of the women. BMI and fat mass were positively skewed. For both variables, log transformations successfully removed the skew. Analyses were thus conducted on these transformed variables. As men are on average taller and have more lean body mass than women (in the current sample, men were 14.7 cm taller, $t(116)=12.08$, $p<.001$, and had 11.9% less body fat, $t(103.2)=9.46$, $p<.001$), height, muscle mass and (log-transformed) fat mass were z-score standardized within sex.

12.1.3. Strength measurements

Two measures of upper body strength were assessed with a hydraulic hand dynamometer (Jamar 5030J1). Handgrip strength was measured following a standard testing protocol three times on the left and the right side with the handle adjusted to a position recommended for testing both men and women (Innes, 1999; Trampisch, Franke, Jedamzik, Hinrichs & Platen, 2012). Participants were tested seated, with their feet flat on the floor, the elbow flexed at a 90° angle with the arm not touching the side of the body, and the forearm in a neutral position. They were instructed to squeeze the handle as hard as they could in a slow, sustained squeeze. The highest grip strength readings from the left and right hand were averaged (Gallup et al., 2007). To measure inverted grip strength or chest strength, subjects were instructed to hold the dynamometer in front of their chest with two hands and compress inwards (Sell et al., 2009; Simmons & Roney, 2011). Again, this procedure was repeated three times. Maximum grip strength and maximum chest strength were separately z-scored within each sex and averaged to produce a composite score of actual strength (Cronbach's α=0.81).

12.2. Identifying anthropometric variables that are predictive of strength

As a first step, zero-order correlations of BMI and height (Study 3.2) and muscle and fat mass (controlling for height, Study 3.3) with the strength composite score were calculated to establish whether or not the measured traits were significantly related to actual strength. Literature suggests that the association of BMI, muscle and fat mass might be sex-specific. Thus, a general linear model was used to test for interactions of sex and height/BMI (Study 3.2), and sex and height/muscle/fat (Study 3.3) in predicting actual strength. If any of the anthropometric traits was found to interact with sex, separate multiple regression analyses were conducted for men and women. Diagnostic regression plots were used to check for normality of residuals, homoscedasticity and outliers. Multicollinearity was considered to be of no concern if tolerance was greater than .10, and the variance inflation factor was less than 3.5.

For one of the women, strength could only be measured for one arm due to an injury; her strength measurements were thus excluded from the analysis. One of the male participants was more than three standard deviations away from the mean height (z-score of 3.1) and was therefore excluded from any analyses involving height.

12.3. Computing, validating and visualizing morphological scores based on group differences

Face-morphological scores were computed as described in Sections 5.2 and 8.2.1.2. Due to the sexual dimorphism in body composition and build, face-morphological scores were separately calculated for men and women. Zero-order correlations of each face score and the variable it was based on were used to test whether face scores captured shape variation associated with the variable of interest (i.e. height, BMI, body fat and muscle mass). In addition, face scores were correlated with each other to test for the independence of face dimensions. All p-values reported are two-tailed.

12.4. Face ratings

12.4.1. Participants

Twenty-seven female and 33 male participants (M_{age} = 35.7±10.1 years, range 22–63) were recruited from the United States of America through Amazon MTurk (Buhrmester et al., 2011). Participants were paid $2.00 each.

12.4.2. Procedure

Prior to the rating, participants were presented with static 2D frontal images of all face models to provide an overview of stimulus variability. The 3D face stimuli were then presented in randomized order, 'bobbing' in a sinusoidal manner from left to right and up and down. For each face, participants were asked "Compared to other men/women his/her age, how physically strong is this person?" Ratings were given on a slider scale beneath each image that ranged from 1-*Very weak* to 100-*Very strong* (numerical values were not visible to participants). Stimuli were presented individually against a black background and remained visible until a rating was made. Female and male faces were presented in two separate blocks; the order of blocks was randomized.

Ratings of strength were z-scored within raters and stimulus sex to account for potential differences in scale use. Ratings were then averaged across participants for each face. Reliability of ratings was calculated using the *R* package *irr* (Gamer, Lemon, Fellows & Singh, 2012; R Development Core Team, 2015). Reliability among raters was high for the average measure (Cronbach's α=.92, 95 % CI [0.90, 0.94]). We note that the intra-class correlation coefficient for the single raters was much lower, though significantly different from 0 (ICC=.16, 95% CI [0.13, 0.20]).

13. Study 3.1: Is strength accurately perceived from 3D faces?

As previous studies were based on 2D colour photographs, the aim of Study 3.1 was to test whether strength can be perceived accurately from colour- and texture-standardized 3D faces. A general linear model was used to test the predictive value of actual strength on ratings of perceived strength, and to test for an effect of stimulus sex. In addition, composite images of faces scoring low and high on actual and perceived strength were created to visualize differences and similarities in face shape associated with actual and perceived strength.

13.1. Results

Actual strength was found to have a significant main effect on perceived strength ($F(1,113)=4.03$, $p=.047$, $\eta_p^2=.034$). Neither the main effect of sex, nor the interaction between sex and actual strength reached significance (both $F(1,113)\leq 0.19$, $p\geq .666$, $\eta_p^2 \leq .002$). Figure 13.1.1 shows the association of actual and perceived strength across both men and women.

Figure 13.1.2 visualizes the face shape associated with actual and perceived strength for men and women.[1] Facial images of the 10 individuals with lowest and highest actual and perceived strength were separately averaged for men and women, resulting in 8 prototypes (2 types of strength [actual, perceived] x 2 levels of strength [low, high] x 2 sexes [male, female], see Table 13.1). The difference in strength between corresponding low and high strength prototypes was calculated and translated into units of standard deviation observed for the respective variable. *Morphanalyser 2.4* was then used to add and subtract the difference between low and high strength prototypes equivalent to ±5 SD of actual and perceived strength to the mean male and female face shape.

For men, shape changes from low to high actual strength were subtle—high strength was associated with a slightly higher forehead, more widely spaced eyebrows and eyes, more pronounced cheekbones (greater bizygomatic width), a longer midface, a wider mouth and, from a frontal view, narrower mandible. For women, high strength was associated with a shorter and rounder face. Compared to women with low strength, women with high strength

[1] Please see `http://perceptionlab.com/~Iris/SI.html` for animated views.

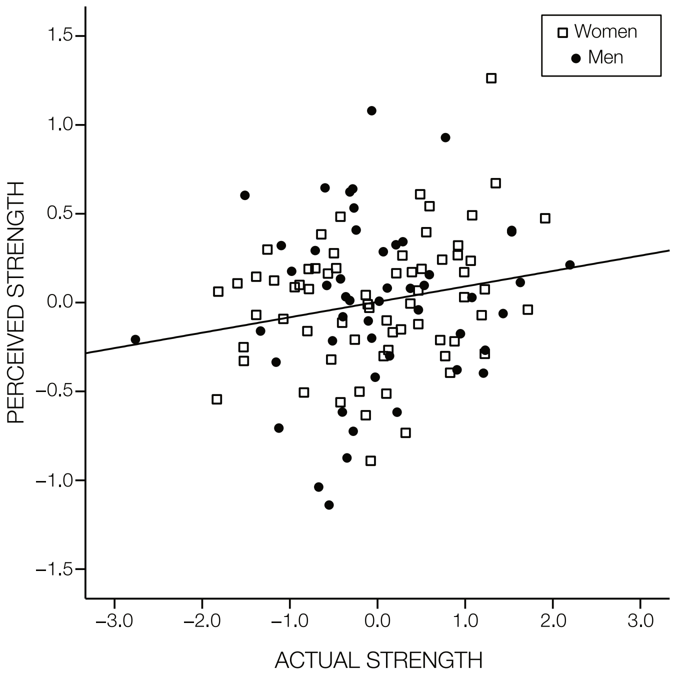


Figure 13.1.1.: Actual strength (a composite of z-scored handgrip and chest strength) was weakly related to perceived strength (average of z-scored ratings, see main text; R^2=.04). The black line represents the best fit regression line for combined male and female face data. Ratings of men (black circles) and women's strength (open squares) did not differ in their accuracy.

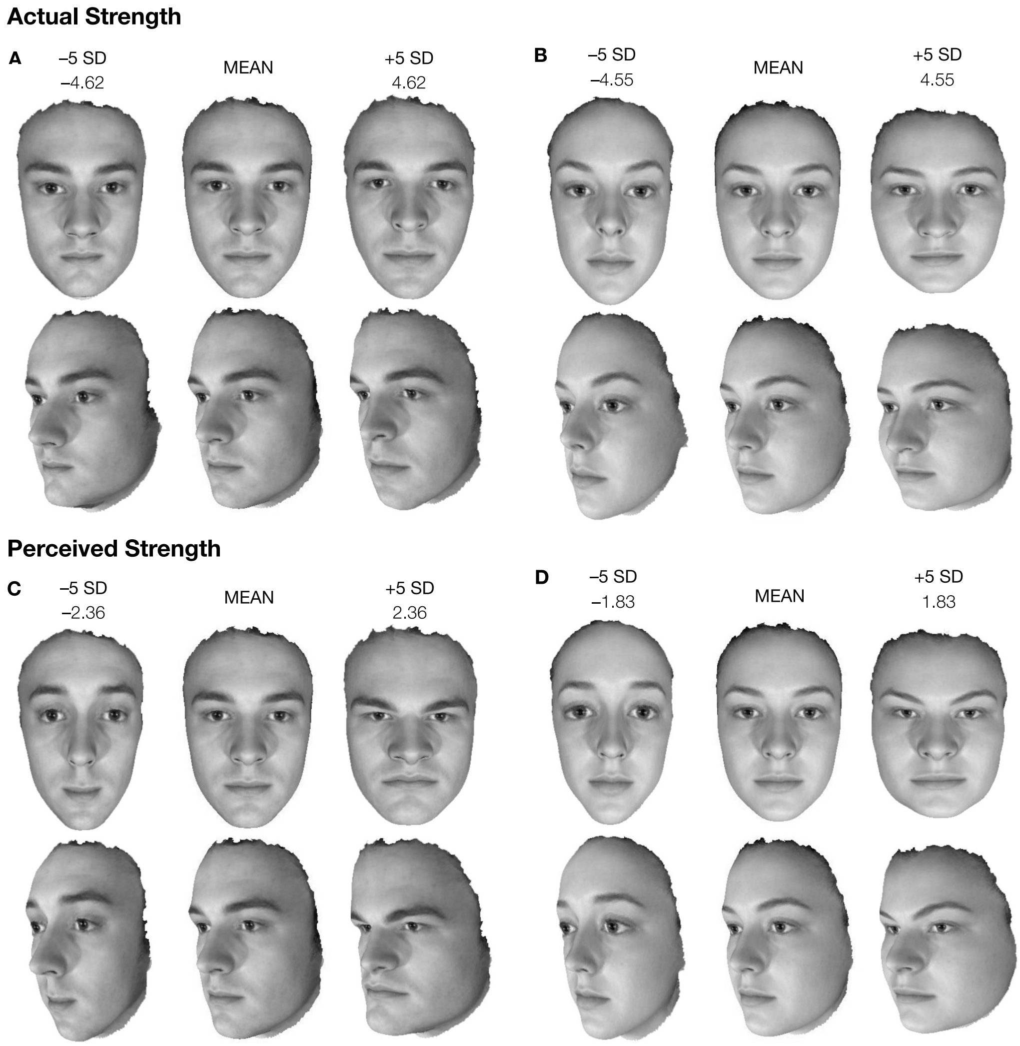


Figure 13.1.2.: Face shape associated with actual (top row) and perceived strength (bottom row). Visualizations reflect face shape associated with ±5 SD of actual and perceived strength in men and women, based on the difference in face shape between the 10 men (A) and women (B) scoring lowest and highest on actual strength, and the 10 men (C) and women (D) scoring lowest and highest on perceived strength (see Table 13.1). Please note that the transform amount of ±5 SD was chosen to increase the salience of changes and goes beyond what would be observed in natural faces.

Table 13.1.: Actual and perceived strength in the tested sample. The difference in shape between composite head models of the 10 men and women scoring lowest and highest on perceived and actual strength served to visualize face shape associated with strength. Due to the missing strength measurement for one participant, female sample size was 67 for actual strength but 68 for perceived strength. Values are given as *M(SD)*. Actual strength is the average of z-scored measure of handgrip and chest strength. Perceived strength was rated on a slider scale from 1-100, z-scored within raters and stimulus sex, and then averaged for each face (see main text). Significant differences ($p<.05$) between low and high prototypes are indicated in bold.

	Sex	Prototype	*N*	Actual Strength	Perceived Strength
Actual Strength	Men	Mean	50	0.00 (0.92)	0.00 (0.47)
		Low	10	**−0.92** (0.35)	0.04 (0.50)
		High	10	**1.33** (0.43)	0.11 (0.40)
	Women	Mean		0.00 (0.91)	0.00 (0.37)
		Low	10	**−1.32** (0.37)	−.09 (0.28)
		High	10	**1.13** (0.32)	.21 (0.45)
Perceived Strength	Men	Mean	50	0.00 (0.92)	0.00 (0.47)
		Low	10	−0.11 (0.72)	**−0.65** (0.29)
		High	10	−0.08 (0.81)	**0.62** (0.23)
	Women	Mean		0.00 (0.91)	0.00 (0.37)
		Low	10	**−0.39** (0.82)	**−0.54** (0.18)
		High	10	**0.55** (0.99)	**0.54** (0.28)

had a shorter forehead, lower brow height and smaller, deeper-set eyes, a shorter midface, a nose that was wider at the level of the nostrils, a wider mouth with thinner lips, a shorter and wider chin, and a wider and more angular mandible.

Perceived strength showed similar facial correlates in men and women. Both men and women's faces that were perceived as stronger had shorter and rounder faces than faces that were perceived as weak. Their foreheads were wider and from a lateral view less bulbous, had a lower brow height, smaller and deeper-set eyes, a shorter midface, a shorter nose (decreased distance between *nasion* and *subnasale*) with a broader bridge and a greater width at the level of the nostrils, a wider mouth, a wider chin and a wider mandible. Men that were perceived as stronger also had a longer and, from a lateral, view more forwardly protruding chin.

13.2. Discussion

In contrast to Sell et al. (2009) and Toscano et al. (2014), we found only a weak relationship between actual and perceived strength. Further, we found no evidence of strength being more accurately perceived from men's as compared to women's faces. Several methodological differences might partly account for these differences in findings. First, the current study used 3D heads, all of which were rendered with the same average skin texture, while Sell et al. (2009) used color 2D photographs. Despite the fact that 3D stimuli likely provide a more comprehensive impression of overall face shape, using a standardized skin texture may conceal shape information that is usually gained through shadows on the face. Second, our stimulus sample size was about half the size of that of Sell et al. (2009). It is therefore likely that actual strength in our study did not vary as much as in Sell et al. (2009) and thus made it harder to detect differences in true strength. Third, Sell et al. (2009) included a self-report measure of strength in their composite measure of actual strength, while we focused on whether the perception of strength is linked to physical predictors of strength.

In accordance with the statistical analysis, visualizing the face shape associated with actual and perceived strength showed similarities between actual and perceived strength in women's but not necessarily men's faces. Women who are stronger, and look stronger, were found to have a rounder face, smaller, deeper-set eyes and lower eyebrows, a shorter and wider nose, and the same facial traits were observed to be associated with perceived strength in men, in line with findings by Toscano et al. (2014). Men's actual strength was linked to only subtle variation in face shape; most notably, and in line with Windhager et al. (2011), a widening between eyebrows and a widening between eyes was observed, as well as an increased bizygomatic width, a wider mouth and a narrower mandible. In contrast to Windhager et al. (2011), male handgrip strength in the current sample was not linked to thinner and higher eyebrows, a shorter nose, thinner lips or a shorter midface.

14. Study 3.2: Do facial cues to BMI and height predict strength?

The aim of Study 3.2 was to test whether perceptions of strength can be linked to face shape associated with BMI and height, two physical characteristics that have been previously found to be predictive of (handgrip) strength. We tested (1) whether BMI and height were related to strength in the current sample, (2) derived face-morphological correlates of BMI and height, and (3) finally tested whether these face scores predict the perception of strength.

14.1. Results

14.1.1. Are BMI and height predictive of strength?

The composite score of actual strength was found to be positively correlated with BMI ($r(117)=.35$, $p<.001$) and height ($r(116)=.22$, $p=.019$; see Table 14.1 for an overview of zero-order correlations of strength and anthropometric measurements). A general linear model (between-subjects factor: stimulus sex [male, female]; covariates: height and BMI) showed no significant interaction between sex and BMI or height in predicting actual strength, nor a main effect of sex (all $F(1,110)\leq 0.50$, $p\geq .479$, $\eta_p^2 \leq .005$). The model was re-run omitting the interaction terms. Effects of BMI ($\beta=0.36$, $t=4.14$, $p<.001$) and height ($\beta=0.22$, $t=2.51$, $p=.013$) on actual strength were significant, while the effect of sex was not ($\beta=0.01$, $t=0.07$, $p=.946$; adj $R^2=.15$, $F(3,112)=7.90$, $p<.001$).

14.1.2. Computing and validating morphological scores of BMI and height

Average values for each PC were separately calculated for men and women with low and high BMI, as well as short and tall men and women. Faces in the low and high groups were matched so that low and high BMI groups did not differ in height, and those in the low and high height groups did not differ in BMI (all $t(18)\leq 0.78$, all $p\geq .454$; see Table 14.2). The difference vectors from low to high height and low to high BMI were used to assign scores to each face on the facial correlates of height and BMI, respectively.

Table 14.1.: Correlation of actual and perceived strength and anthropometric variables: Actual strength (composite of z-scored handgrip and chest strength), perceived strength (average rating derived from z-scores), body mass index (BMI, $\ln(\text{kg m}^{-2})$), height (z(cm)), muscle mass (z(kg)), and fat mass (z(ln(kg)). Values are given as Pearson's *r(N)*.

Group		Actual strength	Perceived strength	BMI	Height	Muscle mass
Overall	Perceived strength	.19* (117)				
	BMI	.35*** (117)	.36*** (118)			
	Height	.22* (116)	.10 (117)	.00 (117)		
	Muscle mass	.49*** (107)	.36*** (108)	.74*** (107)	.52*** (107)	
	Fat mass	.25* (107)	.34*** (108)	.84*** (108)	.29** (107)	.65*** (108)
Men	Perceived strength	.13 (50)				
	BMI	.27 (50)	.41** (50)			
	Height	.19 (49)	.15 (49)	.15 (49)		
	Muscle mass	.50*** (49)	.39** (50)	.73*** (50)	.68*** (49)	
	Fat mass	.06 (50)	.29* (50)	.82*** (50)	.28* (49)	.55*** (50)
Women	Perceived strength	.26* (67)				
	BMI	.40** (67)	.34** (68)			
	Height	.24 (67)	.06 (68)	-.07 (68)		
	Muscle mass	.49*** (57)	.34** (58)	.77*** (58)	.41** (58)	
	Fat mass	.42** (57)	.40** (58)	.87*** (58)	.30* (58)	.74*** (58)

*$p \le .05$, **$p \le .01$, ***$p \le .001$

Table 14.2.: Computing facial correlates of BMI and height (Study 3.2). Face-morphological scores of BMI and height were based on the differences in face shape between men/women with low and high BMI, and differences in face shape of short and tall men/women, respectively. Values are given as *M(SD)*; signficant differences (*p*<.05) between low and high prototypes are indicated in bold.

	Sex	Prototype	*N*	BMI		Height
				kg m^{-2}	ln(kg m^{-2})	cm
BMI	Men	Mean	50	22.7 (2.98)	3.11 (0.13)	181.3 (6.81)
		Low	10	**19.5** (0.94)	**2.97** (0.05)	181.6 (7.58)
		High	10	**26.9** (2.92)	**3.29** (0.10)	182.1 (7.59)
	Women	Mean	68	22.3 (4.11)	3.10 (0.17)	166.6 (6.35)
		Low	10	**17.9** (1.41)	**2.89** (0.08)	167.5 (5.13)
		High	10	**28.7** (2.83)	**3.35** (0.10)	167.7 (4.55)
Height	Men	Mean	50	22.7 (2.98)	3.11 (0.13)	181.3 (6.81)
		Low	10	21.4 (2.78)	3.06 (0.13)	**161.7** (2.67)
		High	10	22.3 (2.59)	3.10 (0.12)	**173.8** (2.89)
	Women	Mean	68	22.3 (4.11)	3.10 (0.17)	166.6 (6.35)
		Low	10	20.7 (2.17)	3.02 (1.00)	**180.8** (4.38)
		High	10	20.4 (2.55)	3.00 (0.13)	**190.6** (4.03)

Face-morphological BMI scores correlated with actual BMI ($r(118)$=.59, p<.001), but not height ($r(117)$=.05, p=.565). Face-morphological height scores correlated with actual height ($r(117)$=.38, p<.001), but not BMI ($r(118)$=.09, p=.323). BMI and height scores were not significantly correlated with each other ($r(118)$=−.10, p=.297). Figures 14.1.1 and 14.1.2 visualize changes in face shape along the BMI and height vector, respectively.[1]

In both men and women, high BMI was associated with a wider, rounder face, smaller eyes, more closely set eyebrows, a narrower nose bridge and greater width at the height of the nostrils, chubbier cheeks (especially in women), wider but less full lips, and a shorter chin. Being taller was in both men and women associated with a more elongated face shape, lower and more closely set eyebrows, a longer chin and a narrower-angled jaw (shorter distance between *gonion* and *pogonion*). In men, being taller was also associated with a larger nose (longer, wider and more curved bridge, wider at the level of the nostrils) and fuller lips, while in women being tall was associated with a shorter, more upward pointing nose, less chubby cheeks, an increased distance between nose and upper lip (philtrum height) and a narrower chin.

[1]Please see `http://perceptionlab.com/~Iris/SI.html` for animated views.

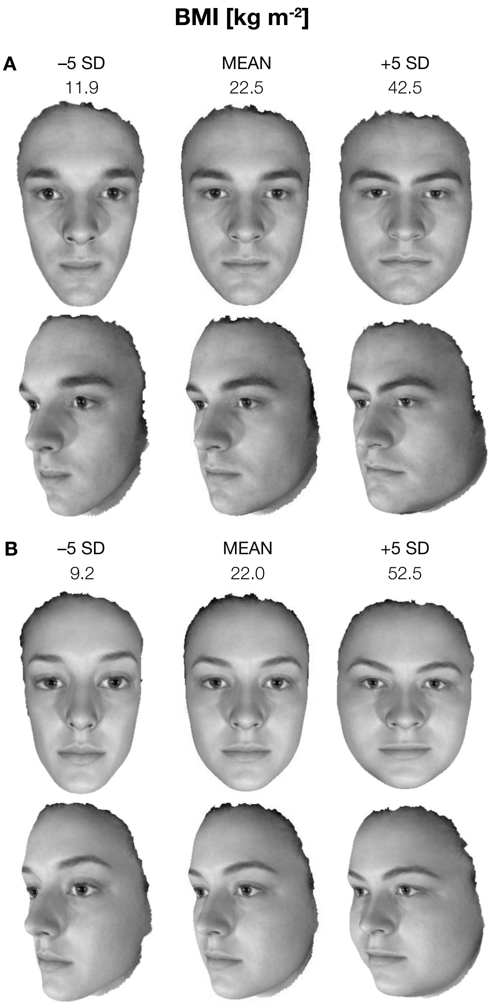


Figure 14.1.1.: Face shape associated with body mass index (BMI) in men (A) and women (B). Faces were manipulated to reflect face shape associated with the sample mean BMI ±5 SD based on the difference in face shape of the low and high BMI prototypes described in Table 14.2. Note that while calculations were based on the log-transformed variable, for the figure numerical values are given on the original scale (kg m^{-2}).

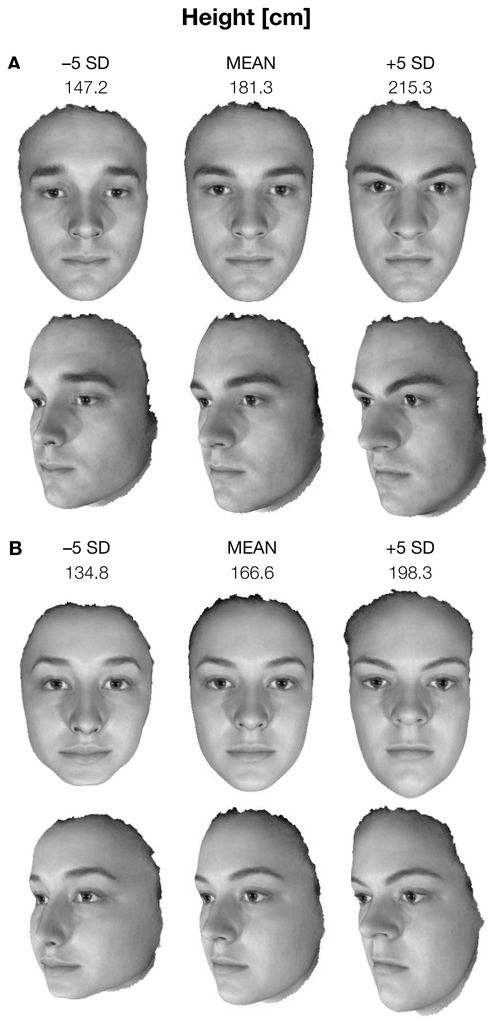


Figure 14.1.2.: Face shape associated with height in men (A) and women (B) based on the difference in face shape of the short and tall prototypes described in Table 14.2.

14.1.3. Do facial correlates of height and BMI predict perceived strength?

The face-morphological height scores were neither related to actual ($r(117)$=.01, p=.943) nor perceived strength ($r(118)$=–.06, p=.531). The face-morphological BMI scores were found to be weakly correlated with actual strength ($r(117)$=.18, p=.054), and strongly correlated with perceived strength ($r(118)$=.53, p<.001).

A general linear model (between-subjects factor: stimulus sex [male, female]; covariates: face-morphological height and BMI scores) showed no main effect of stimulus sex ($F(1,112)$=0.02, p=.897, η_p^2<.001), and no significant interaction between stimulus sex and BMI scores or height scores (both $F(1,112)\leq1.54$, $p\geq.217$, $\eta_p^2\leq.014$). The model was re-run omitting the interaction terms. Again, a significant effect of BMI scores on perceived strength was found (β=0.54, t=6.74, p<.001), while height scores and sex were not predictive of perceived strength (both $\beta\leq.06$, $t\leq0.70$, $p\geq.485$; adj R^2=.27, $F(3,114)$=15.34, p<.001).

To test whether facial correlates of BMI mediated the effect of actual on perceived strength, the *SPSS* plugin *PROCESS* was used (Hayes, 2012). Actual strength was entered as the independent variable, perceived strength as the outcome variable and the face-morphological BMI scores as the mediating variable. Bias-corrected confidence intervals for indirect effects were calculated through 5000 bootstrap samples. Figure 14.1.3 depicts the tested model and results. The completely standardized indirect effect of actual strength on perceived strength (i.e. the mediation effect through the BMI score) was found to be significant (β=0.09, Bootstrap SE=0.05, 95% CI [0.01, 0.21]). The initial significant direct effect of actual on perceived strength was no longer significant (controlling for BMI scores β=0.09, p=.245), confirming the mediation role of facial correlates of BMI in the accuracy of strength perception from faces.

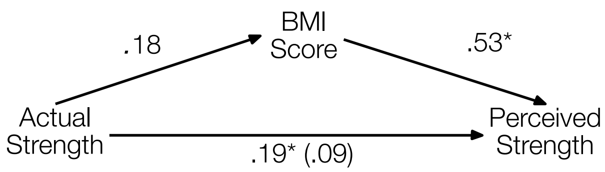


Figure 14.1.3.: Model testing whether the effect of actual strength on perceived strength was mediated by facial cues to BMI (BMI score). Path weights show standardized regression coefficients. The standardized regression coefficient between actual and perceived strength controlling for facial cues to BMI is in parentheses. *p <.05

14.2. Discussion

In line with previous literature, both actual BMI and body height were found to positively predict strength in the current sample. Based on the difference in the average face shape of men and women scoring low and high on these variables, face-morphological scores of BMI and height were computed. The resulting face scores were related to actual BMI and height, but only BMI scores were also related to actual strength. Finally, the BMI score was found to be a strong predictor of perceived strength, and was mediating the effect of actual strength on perceived strength. Thus, the facial correlates of size (BMI) seem responsible for the accuracy in perceptual judgments of strength from 3D face shape in our sample. In line with previous findings, a high BMI was found to be associated with a wider and rounder (mid-) face (e.g., Coetzee et al., 2010), as well as lower and more closely set eyebrows, smaller and deeper-set eyes, wider nose at the level of the nostrils, wider (but not fuller) lips and a shorter lower face (Windhager, Patocka & Schaefer, 2013; Windhager et al., 2011). All of these traits were also found to be associated with perceived strength in Study 3.1. Analyses showed no significant differences in the tested relationships between men and women, suggesting that facial correlates of BMI explain a significant and similar amount of variance in strength perceived from men and women's faces.

15. Study 3.3: Do facial cues to muscle and body fat predict strength?

As BMI might be an inferior indicator of actual strength compared to underlying body composition, Study 3.3 tested for the contribution of facial correlates of muscle and fat mass to perceptions of strength. We (1) tested whether muscle and fat mass were related to actual strength in the current sample, (2) derived face-morphological correlates of muscle and fat and (3) linked them to perceptions of strength.

15.1. Results

15.1.1. Are muscle and fat mass predictive of strength?

The composite score of handgrip and chest strength was found to be positively correlated with muscle mass ($r(107)=.49$, $p\leq.001$) and fat mass ($r(107)=.25$, $p=.011$; see Table 14.1). A general linear model [between-subjects factor: stimulus sex (male, female); covariates: height, muscle and fat mass] showed no significant interaction between stimulus sex and height or muscle mass in predicting actual strength (both $F(1,98)\leq2.25$, $p\geq.137$, $\eta_p^2\leq.022$), but a trend towards an interaction between sex and fat mass ($F(1,98)=3.77$, $p=.055$, $\eta_p^2=.037$). Thus, separate linear models predicting actual strength using the simultaneously entered covariates, muscle mass, fat mass and height, were run for men and women.

For men, actual strength was found to be significantly and positively predicted by muscle mass ($\beta=0.81$, $t=4.11$, $p<.001$) and negatively by fat mass ($\beta=-0.35$, $t=-2.32$, $p=.025$). Height was not significantly related to actual strength ($\beta=-0.26$, $t=-1.49$, $p=.142$; adj $R^2=.25$, $F(3,45)=6.44$, $p=.001$). For women, actual strength again was found to be positively predicted by muscle mass ($\beta=0.38$, $t=2.01$, $p=.050$), but neither height nor fat mass were related to actual strength (both ($\beta\leq0.12$, $t\leq.65$, $p\geq.521$; adj $R^2=.20$, $F(3,53)=5.63$, $p=.002$).

15.1.2. Computing and validating morphological scores of muscle and fat mass

As in Study 3.1, average PC scores were calculated for men and women with low and high absolute muscle mass, as well as men and women with low and high absolute fat mass. Faces in the low and high muscle group were matched so they did not differ in fat mass or height; likewise, faces in the low and high fat group were matched so they did not differ in muscle mass or height (all p≥.461; see Table 15.1). The difference vectors from low to high fat mass and muscle mass were used to assign scores to each face on the facial correlates of fat and muscle, respectively.

Table 15.1.: Computing facial correlates of muscle and fat mass (Study 3.3). Face-morphological scores of muscle and fat were based on the differences in face shape between men/women with low and high muscle/fat mass. Values are given as *M(SD)*; signficant differences (*p*<.05) between low and high prototypes are indicated in bold.

	Sex	Prototype	*N*	Muscle mass kg	Fat mass kg	Fat mass ln(kg)	Height cm	BMI kg m^{-2}
Muscle	Men	Mean	50	61.8 (7.28)	9.7 (5.59)	2.12 (0.58)	181.3 (6.81)	22.7 (2.98)
		Low	10	**57.4** (3.66)	7.9 (4.12)	1.96 (0.50)	181.7 (6.61)	**20.8** (2.02)
		High	10	**65.4** (4.25)	8.5 (5.34)	1.95 (0.66)	180.3 (5.48)	**23.7** (1.62)
	Women	Mean	58	43.4 (4.02)	15.7 (8.00)	2.64 (0.47)	166.6 (6.35)	22.3 (4.11)
		Low	10	**40.5** (2.41)	13.4 (4.15)	2.55 (0.33)	167.2 (3.80)	**20.0** (1.74)
		High	10	**46.6** (3.28)	15.3 (6.96)	2.63 (0.47)	167.0 (4.87)	**23.0** (3.00)
Fat	Men	Mean	50	61.8 (7.28)	9.7 (5.59)	2.12 (0.58)	181.3 (6.81)	22.7 (2.98)
		Low	10	61.2 (5.46)	**4.3** (1.24)	**0.41** (0.32)	180.4 (5.34)	**21.1** (1.84)
		High	10	63.1 (5.58)	**15.8** (4.07)	**2.73** (0.23)	181.1 (4.56)	**25.0** (2.04)
	Women	Mean	58	43.4 (4.02)	15.7 (8.00)	2.64 (0.47)	166.6 (6.35)	22.3 (4.11)
		Low	9	42.9 (3.34)	**9.1** (1.84)	**2.19** (0.21)	165.8 (5.43)	**19.7** (1.02)
		High	9	43.8 (2.42)	**21.2** (4.30)	**3.03** (0.20)	167.0 (6.03)	**24.1** (1.21)

In men, face-morphological muscle scores correlated with muscle mass ($r(50)$=.27, p=.055) but not fat mass ($r(50)$=.06, p=.666) or height ($r(49)$=.15, p=.292). Face-morphological fat scores correlated with fat mass ($r(50)$=.39, p=.005) but not muscle mass ($r(50)$=.14, p=.348) or height ($r(49)$=−.09, p=.552). Face-morphological scores of fat and muscle were not significantly correlated with each other ($r(50)$=−.23, p=.110). Figure 15.1.1 visualizes changes in face shape along the muscle and fat vectors in men.[1]

Higher muscle mass was visually associated with a steeper forehead, a longer mid- and lower face, lower and more closely set eyebrows and more prominent brow ridges, smaller, deeper-set eyes, wider lips, and a longer chin. In addition, high muscle mass seemed to be associated with more prominent cheekbones (i.e. a wider and more pronounced zygomatic arch). Higher amount of body fat was associated with a rounder and wider face, lower, more prominent and more closely set eyebrows, smaller eyes, a smaller nose, wider and thinner lips, and a shorter chin.

[1] Please see `http://perceptionlab.com/~Iris/SI.html` for animated views.

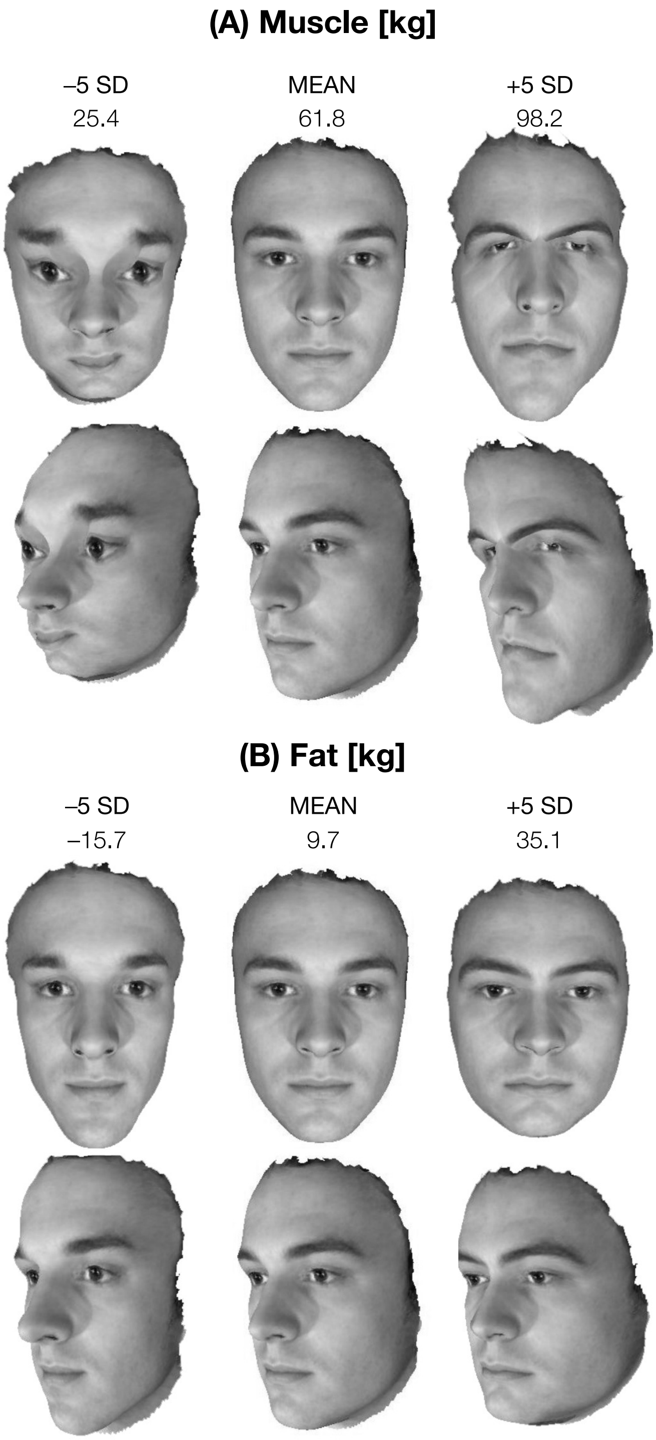


Figure 15.1.1.: Male face shape associated with muscle mass (A) and fat mass (B) based on the difference in face shape between men with low and high muscle and fat mass described in Table 15.1.

In women, no significant association of face-morphological muscle scores and muscle mass was found ($r(58)=.10$, $p=.444$), although this association was significant when controlling for fat mass ($r(55)=.27$, $p=.040$). Muscle scores were not correlated with fat mass ($r(58)=-.11$, $p=.413$) or height ($r(68)=.01$, $p=.971$). Face-morphological fat scores correlated with fat mass ($r(58)=.45$, $p<.001$) and showed a trend to correlate with muscle mass ($r(58)=.23$, $p=.077$) but not height ($r(68)=-.06$, $p=.640$). Face-morphological scores of fat and muscle were significantly correlated with each other ($r(68)=-.52$, $p<.001$), suggesting that we failed to derive separate dimensions of face shape.

15.1.3. Do facial correlates of muscle and fat mass predict perceived strength

In women, face scores of fat and muscle were highly correlated with each other but not necessarily with the variables they were based on, indicating that the face shape associated with muscle and fat could not be satisfactorily separated in women. Thus, the subsequent analysis of the association of muscle- and fat-associated face shape with perceptions of strength was restricted to men's faces.

The facial muscle score showed a trend to relate to actual strength ($r(50)=.25$, $p=.082$) but was not related to perceived strength ($r(50)=.11$, $p=.432$). The fat score was not related to actual strength ($r(50)=-.04$, $p=.770$) but was positively related to perceived strength ($r(50)=.58$, $p<.001$). A general linear model with muscle scores and fat scores as predictors of perceived strength showed significant independent effects of both muscle score ($\beta=0.25$, $t=2.14$, $p=.037$) and fat score ($\beta=0.63$, $t=5.47$, $p<.001$; adj $R^2=.37$, $F(2,47)=15.46$, $p<.001$).

15.2. Discussion

In line with previous literature, the zero-order correlations of actual strength and muscle as well as fat mass showed positive relationships in the current sample. As evidence for an interaction between sex and bodily predictors of strength was found, relationships of fat and muscle were separately investigated for men and women. A multiple linear regression with muscle mass, fat mass and height as predictors of actual strength showed that, for both sexes, muscle mass remained a significant predictor of actual strength when controlling for fat mass and height. In contrast, the relationship of fat and strength differed in the male and female sub-samples when controlling for muscle and height. In women, fat mass was not significantly related to actual strength; in men, fat mass was negatively related to strength. The latter observation is in line with previous findings that fat mass is positively associated with absolute strength, but inversely related to relative strength (i.e. strength per unit muscle mass or strength controlling for muscle mass), although it remains unclear why no such observation was made for women.

As both fat and muscle mass were found to be linked to actual strength, face-morphological scores of muscle and fat mass were derived based on differences in the average face shape of men and women with low and high fat and muscle mass. For men, our results suggested we were successful in describing separate dimensions of face shape associated with muscle and fat mass. The resulting face scores predicted the variable on which they were based (muscle/fat) but were not correlated with the other anthropometric variables (fat and height/ muscle and height), or each other. With regards to women, efforts to separate face shape associated with fat and muscle were unsuccessful. Muscle scores were not associated with any of the anthropometric variables, but highly correlated with fat scores. Fat scores, on the other hand, were related to both fat mass and muscle mass. The difficulties in describing separate dimensions of muscle and fat-associated face shape may reflect the stronger correlation of muscle and fat mass in women compared to men (see Table 14.1). This finding might also reflect a sex difference in sex hormone levels, and in particular testosterone. High testosterone can lead to a greater proportion of lean mass, i.e. a dissociation of fat and muscle, making it easier to separate face shape associated with fat and muscle in men compared to women.

While, for men, we defined two dimensions of face shape change related to distinct body composition components (fat and muscle mass), their perceptual dissociation remains to be shown. Thus, Study 3.4 tested whether face shape associated with fat and muscle would indeed represent two perceivably distinct dimensions in two-alternative forced choice tasks.

The face-morphological fat score was not related to actual strength but was positively related to perceived strength. We note that facial correlates of fat were a stronger predictor of perceived strength than facial correlates of muscle, despite the fact that muscle mass is the stronger predictor of actual strength. Controlling for muscle mass, in men fat mass was negatively correlated with actual strength. Given this negative relationship of fat mass and actual strength in men, these findings are perhaps counterintuitive. They might be better understood by taking into account that in general, increased weight and therefore increased size means higher absolute strength. In line with previous findings (e.g., Lafortuna et al., 2005), zero-order correlations in the current sample showed that fat mass was positively correlated with actual strength overall. Our findings could be interpreted as evidence that observers, above all, use cues to overall size when judging strength from faces. Together, the two face-morphological scores of muscle and fat, derived from absolute muscle and fat mass, both of which were linked to actual strength, explained close to 40% of the variance in ratings of strength.

16. Study 3.4: Are facial cues to muscle and body fat distinct dimensions?

Study 3.3 found that in men (but not women) face shape could be separately related to fat and muscle mass, and two new vectors of male face shape were derived—shape associated with fat mass, and shape associated with muscle mass. Study 3.4 aimed to validate the structural descriptions of fat and muscle mass perceptually. That is, while we derived face shape vectors associated with distinct aspects of body composition—fat and muscle mass—it remained to be established whether the facial shape dimensions would influence perception in distinct and appropriate ways. Study 3.4 thus explored whether the two structural descriptions of muscle and fat mass related to the perception of muscle and fat mass. We designed a two-alternative forced-choice experiment that tested the following two predictions.

1. The defined fat and muscle face shape vectors are perceptually associated with body fat and muscularity.
    a) Manipulating faces towards the shape associated with lower and higher fat mass should affect facial judgments of body fat—'high fat'-faces should be perceived as having more body fat than 'low fat'-faces.
    b) Analogously, manipulating faces towards the shape associated with lower and higher muscle mass should lead 'high muscle'-faces to be perceived as having more muscle than 'low muscle'-faces.
2. Fat- and muscle-associated face shape are separate dimensions.
    a) Manipulating fat-associated face shape should have no effect on perceived muscle mass, while manipulating muscle-associated face shape should have no effect on perceived fat mass.
    b) Comparing high fat- and high muscle-faces, high fat-faces should be perceived as having more body fat than high muscle-faces, while high muscle-faces should be perceived as having more muscle than high fat-faces.

16.1. Methods

16.1.1. Participants

Twenty-five female and 35 male participants (M_{age} = 32.3±8.1 years) were recruited from the United States of America through Amazon MTurk. Participants were paid $2.00 each.

16.1.2. Material

Five male composite faces (each an average of three randomly chosen male faces) were manipulated visually to reflect the face shape associated with low and high levels of muscle and separately fat mass based on the prototypes created in Study 3.3 (see Table 15.1). To visualize the face shape associated with muscle mass, the difference in muscle mass between the low and high muscle prototypes was calculated and translated into standard deviation units (SD) for muscle mass observed in the sample (difference between high and low=7.97 kg equating to 1.09 SD). To visualize the face shape associated with having a muscle mass of 1.50 SD below the mean ('low') and 1.50 SD above the mean ('high'), 1.37 times the difference between low and high prototypes was subtracted from or added to each composite face (as 1.50=1.09*1.37). Analogously, the transform amount equivalent of 1.50 SD of fat mass was subtracted and added from each face to create transforms reflecting the face shape associated with 'low' and 'high' fat mass. Figure 16.1.1 provides an example of the resulting stimuli.

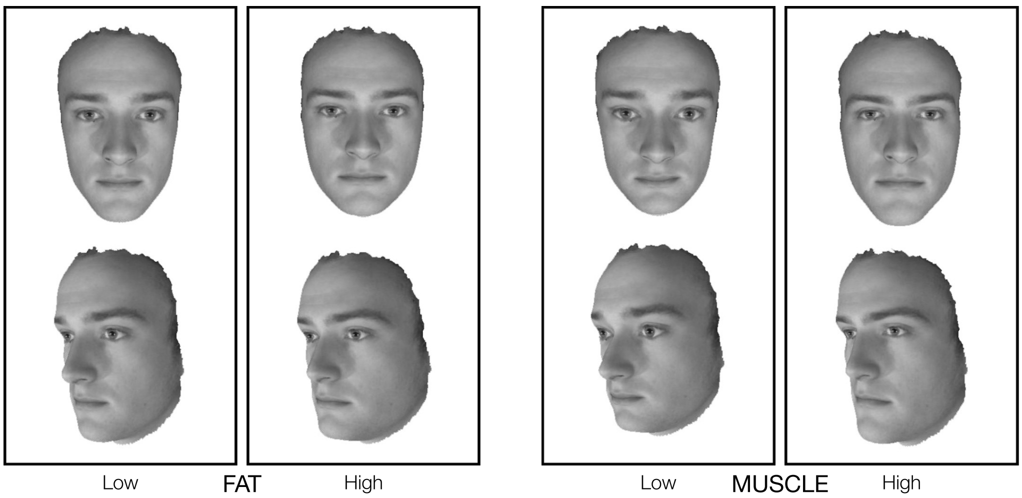


Figure 16.1.1.: Example of stimuli used in validation task. The first and second column show one of the base faces transformed towards the equivalent of 1.5 SD lower (left) and higher (right) fat mass. The third and fourth column show the same base face transformed towards the equivalent of 1.5 SD lower (left) and higher (right) muscle mass.

In total, 20 transforms were generated: five identities x two transform dimensions (muscle and fat) x two transform levels (low and high). These were presented in a two-alternative forced choice task with two different blocks. Participants were asked to choose "which man has more body fat" and "which man has more muscle". In each block, participants were presented with the same 15 face pairs: five pairs of low fat vs high fat, five pairs of low muscle vs high muscle and five pairs of high fat vs high muscle. The order of blocks as well as stimuli within each block was randomized.

16.1.3. Analysis

For each task and stimulus type, the proportion of times a predicted choice was made was calculated. For example, when asked "which man has more body fat", the proportion of trials in which the high fat-face was chosen over the low fat-face was calculated, and separately the proportion of trials in which the high fat-face was chosen over the high muscle-face was calculated. For cross-dimensional choices, such as picking the man with more body fat out of a pair showing low and high muscle transforms, proportions of trials were calculated in which the high transform was chosen over the low transform. As five identities were presented for each stimulus pair combination, the outcome variable could range from 0 to 1, where 0 would indicate that a particular choice was not made once, and 1 would indicate that a particular choice was made for 5 out of 5 identities. Proportions were tested against the null hypothesis of random choice (.50) using one sample t-tests.

16.2. Results

16.2.1. Are the defined face shape vectors perceptually associated with muscle and fat?

A one-sample t-test against chance (.50) showed that when asked "which man has more body fat?", high fat-faces were significantly more often chosen than low fat-faces (.87, $t(59)=17.48$, $p<.001$) and high muscle-faces (.74, $t(59)=7.15$, $p<.001$). When asked "which man has more muscle", high muscle-faces were significantly more often chosen than low muscle-faces (.78, $t(59)=8.638$, $p<.001$) and high fat-faces (.69, $t(59)=5.90$, $p<.001$; see Figure 16.2.1).

16.2.2. Do fat- and muscle-associated face shape describe two separate dimensions?

To test whether fat and muscle vectors described two separate dimensions, cross-dimensional judgments were investigated. For the question, "which man has more muscle mass", no

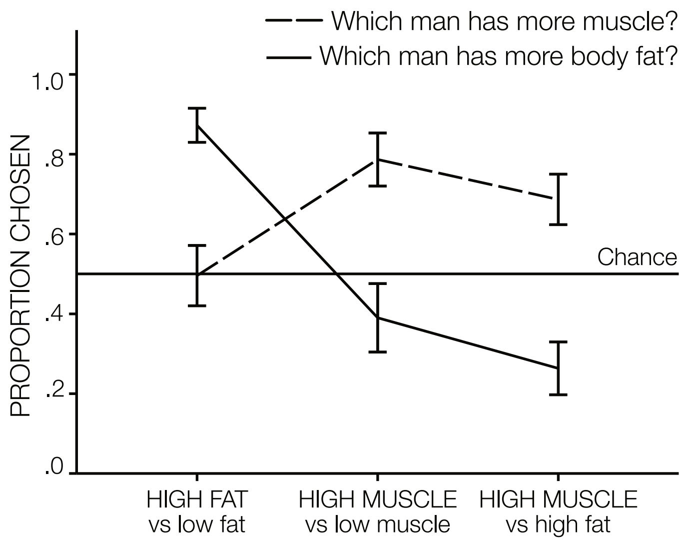


Figure 16.2.1.: Results of the two-alternative forced choice task. Participants were asked to choose in two separate blocks which man out of a pair had more body fat, and which man had more muscle. Participants were presented with three types of stimulus pairs—high fat vs low fat, high muscle vs low muscle and high muscle vs high fat faces. The y-axis gives the proportion with which the capitalized stimulus face was chosen over the lower case-lettered stimulus face. Error bars represent 95% CI.

preference for high or low fat-faces was observed; high fat-faces were chosen as often as low fat-faces (.50, $t(59)=-0.11$, $p=.913$). Contrary to our prediction, when asked "which man has more body fat", participants chose high muscle-faces significantly less often than low muscle-faces (.39, $t(59)=-2.56$, $p=.013$; see Figure 16.2.1).

16.3. Discussion

The fat and muscle vector scores computed in Study 3.3 were found to describe the face shape perceived as being linked to body fat and muscularity, respectively. In addition, we found that these two vectors were perceived as fairly separate dimensions. Men's faces manipulated towards a shape associated with high muscle mass but not high fat mass were perceived as having more muscle. Men's faces manipulated towards a higher fat mass were perceived to have more body fat, although it was found that muscle mass also had an effect on judgments of body fat—face shape associated with lower muscle mass was perceived as having more body fat. These findings suggest that our fat and muscle vectors were successful in describing face shape changes associated with actual fat and muscle mass; they were both correlated with actual fat and muscle mass as well as being perceived as being related to muscle and fat.

17. Discussion

The presented studies investigated whether facial cues to body physique associated with actual strength can account for perceptions of strength from faces. We found that in a set of masked, colour- and texture standardized 3D faces, strength could be assessed with some accuracy. We found BMI as well as body composition (fat and muscle mass) to be linked to both actual strength as well as face shape. The face-morphological correlates of BMI were found to mediate the relationship of actual and perceived strength, explaining about 30% of the variance in perceived strength. In men, further dissecting weight into muscle and fat allowed the separation of two face shape vectors that together explained close to 40% of the variance in perceived strength.

17.1. Facial cues to height and BMI

Body height and BMI were both found to correlate with actual strength. Visualizing the face shape associated with height and BMI showed that a higher BMI was linked to a rounder/wider face (e.g. Coetzee, Greeff, Barrett & Henzi, 2009; Holzleitner et al., 2014), while height was associated with a more elongated face shape (e.g., Holzleitner et al., 2014; Mitteroecker et al., 2013; Re et al., 2013). The computed face-morphological BMI scores were linked to both actual and perceived strength. In contrast, the face-morphological height scores were related to neither actual nor perceived strength. We note that body height was strongly correlated with muscle mass in our sample. The correlation of height and actual strength was no longer significant when controlling for muscle mass, suggesting that it is not height itself that is predictive of strength, but a taller build being associated with a higher amount of lean mass. Visualizing the face shape associated with perception of strength suggested that it is especially the roundness or wideness of a face that drives how strong the face owner looks. We argue that this facial roundness is denoting strength because roundness is a cue to a bulky/heavy body—and on average, heavy means higher strength. We note that this finding might also account for reports that facial width-to-height ratio (fWHR) is linked to perceptions of strength (Zilioli et al., 2014), in line with previous findings that fWHR is correlated with BMI (Coetzee et al., 2010; Lefevre et al., 2012).

17.2. Facial cues to muscle and body fat

Study 3.3 attempted to differentiate facial cues to BMI, or weight, into separate aspects of body composition, muscle and fat mass. Three points are worth noting. First, in men, face shape associated with fat and muscle could be reasonably well separated. Visualizing facial correlates of body fat revealed face shape changes that were closely matched to those associated with BMI. In contrast, the muscle vector revealed overlapping as well as distinct feature changes. For example, high values of BMI/fat and muscle were all associated with more pronounced brow ridges, lower eyebrows and smaller eyes. By contrast, length of mid- and lower face decreased with increasing BMI/fat but increased with increasing muscle mass. Some of the shape changes associated with muscle were reminiscent of shape changes associated with height (such as overall more elongated face shape, e.g., Holzleitner et al., 2014; Mitteroecker et al., 2013; Re et al., 2013) and as outlined earlier, muscle mass increases with increasing height. We note, however, that the prototypes on which muscle vectors were based were matched for height.

It is possible that the muscle vector may be more generally linked to testosterone. Indeed, the muscle-associated face shape revealed characteristics previously described as "masculine" (such as more protruding brow ridges, deeper-set eyes, pronounced cheekbones and a larger jaw). We suggest that effects of testosterone might mediate the perception of strength. Increased muscle mass itself is unlikely to be directly detectable from the face (strength training is unlikely to show in facial musculature). Yet, high levels of testosterone during development will affect body physique/frame size (and hence attainable strength) as well as facial morphology. Observers may use these aspects of facial architecture as cues to body physique and hence strength. As no hormonal measures were taken, this interpretation remains speculative.

Second, efforts to separate fat- and muscle-associated face shape in women were unsuccessful. We suggest this might be due to the stronger correlation of fat and muscle in women than men, which might be linked to the hormonal differences between men and women. In a larger, more varied sample of women it may also be possible to separate face shape associated with muscle and fat.

Third, the three facial features previous linked to perceptions of strength by Toscano et al. (2014) may all be accounted for by the face shape associated with BMI and/or muscle and fat. Our findings show that brow height may be linked to muscularity, nostril width to a heavier body build, and eye size to both weight and muscularity. As both muscularity and BMI were found to be linked to actual strength, our findings may offer an explanation as to why features identified by Toscano et al. (2014) are associated with perceptions of strength.

17.3. Concluding comments

The composite measure of grip and chest strength was only weakly linked to perceived strength in the current sample. Visualizing the face shape associated with perceived strength suggests that, for both male and female faces in the current sample, perceptions of strength were based on similar facial cues. Indeed, Study 3.2 showed that in both sexes a considerable proportion of variance in ratings of perceived strength could be explained by facial cues to BMI or overall mass, such as facial roundness, eyebrows that were narrower and closer together and smaller eyes. Nonetheless, Study 3.3 demonstrated that even more variance in men's perceived strength could be explained by partitioning facial cues to mass into facial cues associated with fat and muscle. Despite a lack of a relationship of actual and perceived strength in men in the current sample, some of the traits that we found to co-vary with perceived strength (such as more pronounced cheekbones and a longer chin) were found to be linked to higher muscle mass, and facial correlates of muscle were found to be linked to both actual as well as perceived strength.

Sell et al. (2009) found that in three out of four tested samples, measured upper-body strength was a better predictor of men's perceived strength than body weight. They concluded that judgments of strength from faces track muscularity rather than overall body size. We interpret our findings slightly differently. We agree that muscularity is a cue to strength, yet we note that overall size may be a more effective perceptual cue to strength. Our study is the first to identify facial correlates of muscularity in 3D face shape. By directly testing for the effect of facial shape correlates of muscle mass as well as fat mass and overall mass (BMI), we find that muscularity is a significant predictor of perceived strength. At the same time, facial correlates of overall body size had a stronger effect on perceptions of strength than facial correlates of muscularity. Indeed, in line with our findings, Sell and colleagues did find that for women and men in their US sample, the effect of body weight was equal to or larger than the effect of actual strength on perceived strength.

Taken together, findings from the current study provide limited support for suggestions that men's face shape evolved as a signal of formidability. Some aspects of men's face shape that seem to influence the perception of strength (such as facial adiposity or muscularity) could be a 'by-product' of a selection pressure for overall greater body size. These aspects of face shape do not need to have or have had automatic signal value; instead their link to physical characteristics (and hence strength) could be learned. Other, and maybe less physique-dependent aspects of facial shape, could have been selected for. For example, a larger and more robust zygomatic arch might result from benefits associated with a larger masseter muscle and greater bite force. Alternatively, greater robusticity might have been beneficial by providing greater resilience to contact violence (Stirrat et al., 2012).

Despite the fact that we found actual strength to be only weakly associated with perceived strength, we have shown that perceptions of strength are likely rooted in facial correlates of physical parameters. Facial correlates of BMI, a rough measure of overall size or bulk, were found to be strongly predictive of perceptions of strength in both men and women. Future studies could further investigate the relationship of sex hormone levels, body composition and facial correlates of body composition. If facial sexual dimorphism is partly mediated by dimorphism in body composition, accounting for these sex differences might allow for a more targeted investigation of sexually selected facial traits.

Part V.

Synthesis and Reflections

Summary

Recent advances in computer graphic and statistical methods have made it possible to visualize global face shape correlates of social judgments. The current thesis used a data-driven approach to investigate face shape correlates and perception of two traits, masculinity and strength, both of which have been previously found to have important effects on mate choice and social perception more generally. The studies presented defined the influences of body physique (height, body mass index, body fat and muscle mass) on facial shape, and via these influences on the perception of masculinity and strength.

Study 1 investigated the face shape correlates of actual and perceived masculinity. In line with previous studies, I found that morphological masculinity explains only limited variance in perceptions of masculinity. This finding begs caution regarding inferences on the adaptiveness of women's preferences for a masculine face shape: perceived masculinity has often been used as a measure in this context yet may not be entirely appropriate. My results show that perceived masculinity is not only driven by sexually dimorphic shape, but also by cues to body height and weight. Men with taller, and heavier bodies were perceived to have more masculine looking faces.

Study 2 investigated women's perception of male attractiveness as a function of masculine face shape. As previously assumed but not explicitly tested, masculinity preferences were found to follow a quadratic relationship: attractiveness increased with increasing masculinity levels, but dropped off at higher levels of masculinity. In addition, I showed that the attractiveness function of masculinity is affected by individual differences in own condition, perceived financial harshness and pathogen disgust. Women who perceived themselves as low in attractiveness or who worried about finance showed a greater aversion to high masculine male faces and/or a greater attraction to more feminine male faces.

Study 3 demonstrated that perception of strength from faces is driven by facial cues to body physique. Individuals with higher body bulk were perceived to be stronger. In men it proved possible to further dissect facial cues to muscle mass and fat mass; cues that both contributed to perception strength. Findings showed that perception reflected cues to absolute strength rather than strength per unit body size.

Theoretical contributions, limitations and future prospects

My findings show the importance of facial cues to height and weight, or overall size, in perceptions of both masculinity and strength. In line with suggested overgeneralization effects in social perception, facial cues to height and weight predicted perceptions of height and weight more accurately than actual height and weight. Likewise, facial cues to size were more strongly linked to perceived than to actual sexual dimorphism and strength.

Future studies on the attractiveness of masculinity and on the effects of facial masculinity on perceived dominance and leadership ability should follow-up on evidence that (masculine) face shape may be comprised of two different dimensions: one related to overall size, one independent of size and likely related to sex hormone influences. Facial cues to height have already been shown to be linked to perceptions of dominance and leadership ability. Findings from the biological literature and this thesis suggest that a more thorough investigation of the face shape changes associated with size scaling and consequent effects on mate choice and social judgments is warranted.

My work is the first to dissociate facial cues to weight into facial cues to fat and muscle mass. While the face shape correlates of fat show resemblance to previous visualisations of BMI, the derived muscle vector shows resemblance to face shape correlates of both height and masculinity. I suggest that this muscle vector might be linked to androgens such as testosterone, and potentially be a visual indicator of somatic reproductive effort: resources allocated to muscle mass, thereby facilitating male competitiveness. Future studies should investigate more thoroughly how testosterone levels and testosterone sensitivity are linked to physical and facial characteristics and how these characteristics in turn affect social judgements.

Previous studies found that facial adiposity in a developed country was negatively linked to perceived attractiveness (Rantala et al., 2013). Yet, my finding that facial cues to weight are linked to perceptions of strength and masculinity in a similar, Western sample[1] would suggest that facial cues to weight, to a certain extent, should have a positive effect on perceived attractiveness.

One possible explanation for this discrepancy might be that Rantala et al. collected attractiveness and adiposity ratings from colour-photographs. Textural and reflectance cues have been suggested to be important cues to current health status, while shape cues such as those to masculinity may reflect developmental stability or long-term condition (Scott et al., 2010; Stephen et al., 2012). Cues to current and long-term condition may be of differential importance in different contexts (e.g., when choosing a short- or long-term partner). Moreover, reflectance and shape associated with the same trait (e.g., sex) have been found to affect preferences differently, with shape masculinity detracting from attractiveness and reflectance cues to masculinity enhancing attractiveness (e.g., Said & Todorov, 2011). It is likely that if both shape and colour information are available, they interact in affecting perceptions and preferences. That is, while a face of high weight when presented with its

[1] In rural (harsh) environments, a higher body mass is often found more attractive than lower body weight, although it is unknown whether or not this reflects attraction to physiques with more fat or muscle (e.g., Batres & Perrett, 2014; Swami & Tovée, 2005).

original texture and colour may be perceived as unhealthy due to hormonal effects of, e.g., cortisol on skin colour, face shape cues to weight or bulk might be perceived as strong and healthy when skin colouration and texture are held constant as in the current work. One of the limitations of the current work, thus, is that it did not investigate how shape cues are integrated with colour cues. I started my investigation with a focus on a more refined analysis of shape, but future studies should extend research to the relative importance of textural and shape information.

Methodological contributions

The current work presented several new approaches that should prove valuable in future investigations. First, we did not only visualize (and manipulate) facial cues driving perceptual judgments, but we also made these shape cues available to statistical modelling and analysis. For example, in Study 1 statistical analysis showed that facial cues to body height and weight seem to contribute more to the perception of masculinity than sexual dimorphism in shape (i.e. the difference between average male and female shape). Likewise, in Study 3 regression analysis showed that facial cues to fat are better predictors of perceived strength than facial cues to muscle despite being weaker predictors of actual strength. Such findings are in line with overgeneralization effects found for other perceptual judgments.

Second, while many studies have worked with 3D shape models of faces, participants in these studies evaluated static 2D frontal images of the 3D face models. That is, many potential perceptual effects of 3D information are discarded in studies with static 2D images. The current thesis made an effort to retain 3D information in perceptual studies. This was done by animating changes in perspective view of the 3D shape. The rational behind this effort was that 3D stimuli confer a higher degree of ecological validity.

The current thesis cannot give a definitive answer as to whether the ecological validity of 3D stimuli is indeed higher. Perceptual judgments from 2D and 3D faces were compared for two traits, masculinity and weight. Due to the discrepancy of morphological and perceived masculinity, masculinity does not lend itself to an investigation of perceptual accuracy; with regards to weight, no differences between 2D and 3D stimuli was found. Yet, it would be premature to conclude that the use of 3D stimuli is not preferable over 2D stimuli. First, more traits should be compared in 2D and 3D. Second, stimuli in the current thesis suffered in their realism by the low-resolution of their texture images and shadows locked to the face, rather than changing with changes in view. With a newer camera system, and imaging software, this could be easily remedied.

Third, prior literature had speculated on the trade off between the potential costs associated with high masculinity (greater infidelity, lower cooperativeness) and the potential benefits (e.g., better heritable immunocompetence, higher status, higher competitiveness). Analysing the relationship between masculinity and attractiveness allowed inferences about the consequences of relative costs and benefits to women dependent on their own condition. The experimental design and analytic methods of Study 2 may be useful in the investigation of a number of traits other than masculinity. Cost and benefit functions of traits could be explored to test for context-specificity within subjects or for cultural differences between subjects. For example, different levels of trustworthiness or dominance might be of different value in different political contexts (e.g., outgroup vs ingroup conflict).

Conclusion

The work of the current thesis is only one step towards a more refined investigation of the basis and functional significance of social perceptions. Previous research suggested that perceptions of dominance are adaptive because they reflect facial cues to strength. The current thesis demonstrated that facial cues used in the evaluation of masculinity and strength are linked to bodily characteristics associated with sex differences and actual strength, namely height, weight, muscularity and adiposity. The thesis findings therefore support the hypothesis that perceptions have an adaptive origin. In a next step, face shape correlates of determinants of physical prowess should be linked to judgments of dominance and attractiveness, and integrated with social perception.

References

Alley, T. R. & Cunningham, M. R. (1991). Averaged faces are attractive, but very attractive faces are not average. *Psychological Science*, *2*(2), 123–125. doi: 10.1111/j.1467-9280.1991.tb00113.x

Alonso-Alvarez, C., Bertrand, S. & Sorci, G. (2007). Energetic reserves, leptin and testosterone: a refinement of the immunocompetence handicap hypothesis. *Biology letters*, *3*(3), 271–274. doi: 10.1098/rsbl.2007.0020

Alvergne, A. & Lummaa, V. (2010). Does the contraceptive pill alter mate choice in humans? *Trends in Ecology and Evolution*, *25*(3), 171–179. doi: 10.1016/j.tree.2009.08.003

Andersson, M. (1994). *Sexual Selection* (J. R. Krebs & T. Clutton-Brock, Eds.). Princeton, New Jersey: Princeton University Press.

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

Balogun, J. A., Akinloye, A. A. & Adenlola, S. A. (1991). Grip strength as a function of age, height, body weight and quetelet index. *Physiotherapy Theory and Practice*, *7*(2), 111–119. doi: 10.3109/09593989109106961

Bar, M., Neta, M. & Linz, H. (2006). Very first impressions. *Emotion (Washington, D.C.)*, *6*(2), 269–278. doi: 10.1037/1528-3542.6.2.269

Bateman, A. J. (1948). Intra-sexual selection in Drosophila. *Heredity*, *2*(3), 349–368.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4.* Retrieved from `http://cran.r-project.org/package=lme4`

Batres, C., Jones, B. C., Feinberg, D. R., DeBruine, L. M. & Perrett, D. I. (n.d.). Variation in sexual orientation level amongst heterosexual women predicts facial masculinity preferences in male faces. *Unpublished Manuscript*.

Batres, C. & Perrett, D. I. (2014). The influence of the digital divide on face preferences in El Salvador: people without internet access prefer more feminine men, more masculine women, and women with higher adiposity. *PLoS ONE*, *9*(7), e100966. doi: 10.1371/journal.pone.0100966

Bellis, M. A. & Baker, R. R. (1990). Do females promote sperm competition? Data for humans. *Animal Behaviour*, *40*(5), 997–999.

Benson, P. J. & Perrett, D. I. (1991, apr). Synthesising continuous-tone caricatures. *Image and Vision Computing*, *9*(2), 123–129. doi: 10.1016/0262-8856(91)90022-H

Berry, D. S. & Wero, J. L. F. (1993). Accuracy in face perception: A view from ecological psychology. *Journal of Personality*, *61*(4), 497–520.

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

Berry, D. S. & Zebrowitz McArthur, L. (1988). What's in a face? Facial maturity and the attribution of legal responsibility. *Personality and Social Psychology Bulletin*, *14*(1), 23–33.

Bhasin, S., Woodhouse, L. & Storer, T. W. (2003). Androgen effects on body composition. *Growth Hormone & Igf Research*, *13*, S63–S71. doi: 10.1016/S1096-6374(03)00058-3

Blanz, V. & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *Siggraph 99 Conference Proceedings*, 187–194.

Blanz, V. & Vetter, T. (2003). Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(9), 1063–1074. doi: 10.1109/TPAMI.2003.1227983

Boonekamp, J. J., Ros, A. H. F. & Verhulst, S. (2008). Immune activation suppresses plasma testosterone level: A meta-analysis. *Biology Letters*, *4*(6), 741–744. doi: 10.1098/rsbl.2008.0347

Booth, A. & Dabbs, J. M. (1993). Testosterone and men's marriages. *Social Forces*, *72*(2), 463–477.

Boothroyd, L. G. & Brewer, G. (2014). Self-reported impulsivity, rather than sociosexuality, predicts women's preferences for masculine features in male faces. *Archives of Sexual Behavior*, *43*(5), 983–988. doi: 10.1007/s10508-013-0204-z

Boothroyd, L. G., Jones, B. C., Burt, D. M., Cornwell, R. E., Little, A. C., Tiddeman, B. P. & Perrett, D. I. (2005). Facial masculinity is related to perceived age but not perceived health. *Evolution and Human Behavior*, *26*(5), 417–431.

Boothroyd, L. G., Jones, B. C., Burt, D. M., DeBruine, L. M. & Perrett, D. I. (2008). Facial correlates of sociosexuality. *Evolution and Human Behavior*, *29*(3), 211–218.

Boothroyd, L. G., Jones, B. C., Burt, D. M. & Perrett, D. I. (2007). Partner characteristics associated with masculinity, health and maturity in male faces. *Personality and Individual Differences*, *43*(5), 1161–1173.

Boothroyd, L. G. & Perrett, D. I. (2008). Father absence, parent-daughter relationships and partner preferences. *Journal of Evolutionary Psychology*, *6*(3), 187–205. doi: 10.1556/JEP.6.2008.3.3

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. doi: 10.1016/j.jrp.2009.03.007

Brand, A. & Bradley, M. T. (2012). More voodoo correlations: when average-based measures inflate correlations. *The Journal of General Psychology*, *139*(4), 260–272. doi: 10.1080/00221309.2012.703711

Brooks, R., Scott, I. M. L., Maklakov, A. A., Kasumovic, M. M., Clark, A. P. & Penton-Voak, I. S. (2011). National income inequality predicts women's preferences for masculinized faces better than health does. *Proc. R. Soc. B*, *278*(1707), 810–812. doi: 10.1098/rspb.2010.0964

Brown, G. R., Laland, K. N. & Mulder, M. B. (2009). Bateman's principles and human sex roles. *Trends in Ecology and Evolution*, *24*(6), 297–304.

Bruce, V., Burton, A. M., Hanna, E., Healey, P., Mason, O., Coombes, A., … Linney, A. (1993). Sex discrimination: how do we tell the difference between male and female faces? *Perception*, *22*(2), 131–152. doi: 10.1068/P220131

Bruce, V. & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327. doi: 10.1111/j.2044-8295.1986.tb02199.x

Buhrmester, M., Kwang, T. & Gosling, S. D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, *6*(1), 3–5. doi: 10.1177/1745691610393980

Burley, N. (1986). Sexual Selection for Aesthetic Traits in Species with Biparental Care. *The American Naturalist*, *127*(4), 415–445. doi: 10.1086/284493

Burnham, T. C., Chapman, J. F., Gray, P. B., McIntyre, M. H., Lipson, S. F. & Ellison, P. T. (2003). Men in committed, romantic relationships have lower testosterone. *Hormones and Behavior*, *44*(2), 119–122.

Burton, A. M., Bruce, V. & Dench, N. (1993). What's the difference between men and women? Evidence from facial measurement. *Perception*, *22*(2), 153–176. doi: 10.1068/P220153

Buss, D. M. (1989). Sex differences in human mate preferences: evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, *12*, 1–49.

Buss, D. M. (2008). *Evolutionary psychology: the new science of the mind* (3rd ed.). Pearson.

Buss, D. M. & Schmitt, D. P. (1993). Sexual strategies theory: an evolutionary perspective on human mating. *Psychological Review*, *100*(2), 204–232.

Buss, D. M. & Schmitt, D. P. (2011). Evolutionary psychology and feminism. *Sex Roles*, *64*(9), 768–787.

Carré, J. M. & McCormick, C. M. (2008). In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proc. R. Soc. B*, *275*(1651), 2651–2656. doi: DOI10.1098/rspb.2008.0873

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.

Cash, T. F., Gillen, B. & Burns, D. S. (1977). Sexism and "beautyism" in personnel consultant decision making. *Journal of Applied Psychology*, *62*(3), 301–310.

Chandrasekaran, B., Ghosh, A., Prasad, C., Krishnan, K. & Chandrasharma, B. (2010). Age and Anthropometric Traits Predict Handgrip Strength in Healthy Normals. *Journal of Hand and Microsurgery*, *2*(2), 58–61. doi: 10.1007/s12593-010-0015-6

Coetzee, V., Chen, J. Y., Perrett, D. I. & Stephen, I. D. (2010). Deciphering faces: quantifiable visual cues to weight. *Perception*, *39*(1), 51–61. doi: 10.1068/P6560

Coetzee, V., Greeff, J. M., Barrett, L. & Henzi, S. P. (2009). Facial-based ethnic recognition: Insights from two closely related but ethnically distinct groups. *South African Journal of Science*, *105*(11-12), 464–466. doi: 10.4102/sajs.v105i11/12.127

Collins, M. A. & Zebrowitz, L. A. (1995). The contributions of appearance to occupational outcomes in civilian and military settings. *Journal of Applied Social Psychology*, *25*(2), 129–163.

Cornwell, R. E. & Perrett, D. I. (2008). Sexy sons and sexy daughters: the influence of parents' facial characteristics on offspring. *Animal Behaviour*, *76*(6), 1843–1853. doi: 10.1016/j.anbehav.2008.07.031

Craw, I. & Cameron, P. (1991). Parameterising images for recognition and reconstruction. In P. Mowforth (Ed.), *Proceedings of the british machine vision conference* (pp. 367–370). Springer London. doi: http://dx.doi.org/10.1007/978-1-4471-1921-0{_}52

Cunningham, M. R. (1986). Measuing the physical in physical attractiveness: quasi-experiments on the sociobiology of female facial beauty. *Journal of Personality and Social Psychology*, *50*(5), 925–935.

Cunningham, M. R., Barbee, A. P. & Pike, C. L. (1990). What do women want? Facialmetric assessment of multiple motives in the perception of male physical attractiveness. *Journal of Personality and Social Psychology*, *59*(1), 61–72.

De Barra, M., DeBruine, L. M., Jones, B. C., Mahmud, Z. H. & Curtis, V. A. (2013). Illness in childhood predicts face preferences in adulthood. *Evolution and Human Behavior*, *34*(6), 384–389. doi: 10.1016/j.evolhumbehav.2013.07.001

DeBruine, L. M., Jones, B. C., Crawford, J. R., Welling, L. L. M. & Little, A. C. (2010). The health of a nation predicts their mate preferences: cross-cultural variation in women's preferences for masculinized male faces. *Proc. R. Soc. B*, *277*(1692), 2405–2410. doi: 10.1098/rspb.2009.2184

DeBruine, L. M., Jones, B. C., Frederick, D. A., Haselton, M. G., Penton-Voak, I. S. & Perrett, D. I. (2010). Evidence for menstrual cycle shifts in women's preferences for masculinity: a response to Harris (in press) "Menstrual cycle and facial preferences reconsidered". *Evolutionary Psychology*, *8*(4), 768–775.

DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., …

Tiddeman, B. P. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proc. R. Soc. B*, *273*(1592), 1355–1360. doi: 10.1098/rspb.2005.3445

DeBruine, L. M., Jones, B. C., Smith, F. G. & Little, A. C. (2010). Are attractive men's faces masculine or feminine? The importance of controlling confounds in face stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(3), 751–758.

DeBruine, L. M., Jones, B. C., Tybur, J. M., Lieberman, D. & Griskevicius, V. (2010). Women's preferences for masculinity in male faces are predicted by pathogen disgust, but not by moral or sexual disgust. *Evolution and Human Behavior*, *31*(1), 69–74.

DeBruine, L. M., Jones, B. C., Watkins, C. D., Roberts, S. C., Little, A. C., Smith, F. G. & Quist, M. C. (2011). Opposite-sex siblings decrease attraction, but not prosocial attributions, to self-resembling opposite-sex faces. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(28), 11710–11714. doi: 10.1073/pnas.1105919108

Deurenberg, P., Yap, M. & van Staveren, W. A. (1998). Body mass index and percent body fat: a metaanalysis among different ethnic groups. *International Journal of Obesity*, *22*(12), 1164–1171.

Dion, K., Walster, E. & Berscheid, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*(3), 285–290.

Dipboye, R. L., Arvey, R. D. & Terpstra, D. E. (1977). Sex and physical attractiveness of raters and applicants as determinants of resume evaluations. *Journal of Applied Psychology*, *62*(3), 288–294. doi: 10.1037/0021-9010.62.3.288

Dipboye, R. L., Fromkin, H. L. & Wiback, K. (1975). Relative importance of applicant sex, attractiveness, and scholastic standing in evaluation of job applicant resumés. *Journal of Applied Psychology*, *60*(1), 39–43.

Dzhelyova, M., Perrett, D. I. & Jentzsch, I. (2012). Temporal dynamics of trustworthiness perception. *Brain Research*, *1435*, 81–90. doi: 10.1016/j.brainres.2011.11.043

Eens, M. & Pinxten, R. (2000). Sex-role reversal in vertebrates: behavioural and endocrinological accounts. *Behavioural Processes*, *51*(1-3), 135–147.

Ellison, P. T. (2003). Energetics and reproductive effort. *American Journal of Human Biology*, *15*(3), 342–351. doi: 10.1002/ajhb.10152

Engell, A. D., Haxby, J. V. & Todorov, A. (2007). Implicit trustworthiness decisions: automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience*, *19*(9), 1508–1519. doi: 10.1162/jocn.2007.19.9.1508

Enlow, D. H. & Hans, M. G. (1996). *Essentials of facial growth*. Philadelphia: W.B. Saunders Company.

Ewing, L., Caulfield, F., Read, A. & Rhodes, G. (2014). Perceived trustworthiness of faces drives trust behaviour in children. *Developmental Science*, *2*, 327–334. doi: 10.1111/desc.12218

Farkas, L. G. (1994). *Anthropometry of the Head and Face* (2nd Editio ed.). New York: Raven

Press.

Feinberg, D. R., DeBruine, L. M., Jones, B. C., Little, a. C., O'Connor, J. J. M. & Tigue, C. C. (2012). Women's self-perceived health and attractiveness predict their male vocal masculinity preferences in different directions across short- and long-term relationship contexts. *Behavioral Ecology and Sociobiology*, *66*(3), 413–418. doi: 10.1007/s00265-011-1287-y

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin*, *111*(2), 304–341.

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.

Fink, B., Neave, N., Manning, J. T. & Grammer, K. (2005). Facial symmetry and the 'big-five' personality factors. *Personality and Individual Differences*, *39*(3), 523–529. doi: 10.1016/j.paid.2005.02.002

Fink, B., Neave, N. & Seydel, H. (2007). Male facial appearance signals physical strength to women. *American Journal of Human Biology*, *19*(1), 82–87.

Fink, B., Thanzami, V., Seydel, H. & Manning, J. T. (2006). Digit ratio and hand-grip strength in German and Mizos men: Cross-cultural evidence for an organizing effect of prenatal testosterone on strength. *American Journal of Human Biology*, *18*(6), 776–782.

Fink, B., Weege, B., Manning, J. T. & Trivers, R. (2014). Body Symmetry and Physical Strength in Human Males. *American Journal of Human Biology*, *26*(5), 697–700. doi: 10.1002/Ajhb.22584

Fink, B., Weege, B., Neave, N., Pham, M. N. & Shackelford, T. K. (2015). Integrating body movement into attractiveness research. *Frontiers in Psychology*, *6*. doi: 10.3389/fpsyg.2015.00220

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford Calarendon Press.

Folstad, I. & Karter, A. J. (1992). Parasites, bright males, and the immunocompetence handicap. *American Naturalist*, *139*(3), 603–622. doi: 10.1086/285346

Forbes, G. B. (1987). *Human Body Composition: Growth, Aging, Nutrition, and Activity*. New York: Springer-Verlag New York Inc.

Forbes, G. B. (1993). The Companionship of Lean and Fat. In K. J. Ellis & J. D. Eastman (Eds.), *Human body composition: In vivo methods, models, and assessment* (pp. 1–14). New York and London: Plenum Press.

Fraccaro, P. J., Feinberg, D. R., DeBruine, L. M., Little, A. C., Watkins, C. D. & Jones, B. C. (2010). Correlated Male Preferences for Femininity in Female Faces and Voices. *Evolutionary Psychology*, *8*(3), 447–461. doi: 10.1.1.174.4476

Gallup, A. C., White, D. D. & Gallup Jr, G. G. (2007). Handgrip strength predicts sexual behavior, body morphology, and aggression in male college students. *Evolution and Human Behavior*, *28*(6), 423–429. doi: 10.1016/j.evolhumbehav.2007.07.001

Gamer, M., Lemon, J., Fellows, I. & Singh, P. (2012). *irr: Various coefficients of interrater reliability and agreement* (.84 ed.). Retrieved from `http://cran.r-project.org/package=irr`

Gangestad, S. W. & Buss, D. M. (1993). Pathogen prevalence and human mate preferences. *Ethology and Sociobiology*, *14*, 89–96.

Gangestad, S. W. & Eaton, M. A. (2013). Toward an integrative perspective on sexual selection and men's masculinity. *Behavioral Ecology*, *24*(3), 594–595. doi: 10.1093/beheco/ars096

Gangestad, S. W. & Simpson, J. A. (2000). The evolution of human mating: Trade-offs and strategic pluralism. *Behavioral and Brain Sciences*, *23*(4), 573–587.

Gangestad, S. W. & Thornhill, R. (1998). Menstrual cycle variation in women's preferences for the scent of symmetrical men. *Proc. R. Soc. B*, *265*(1399), 927–933. doi: 10.1098/rspb.1998.0380

Gangestad, S. W. & Thornhill, R. (2008). Human oestrus. *Proc. R. Soc. B*, *275*(1638), 991–1000. doi: 10.1098/rspb.2007.1425

Gaulin, S. & Boster, J. (1985). Cross-cultural differences in sexual dimorphism. Is there any variance to be explained? *Ethology and Sociobiology*, *6*(4), 219–225.

Gildersleeve, K., DeBruine, L. M., Haselton, M. G., Frederick, D. A., Penton-Voak, I. S., Jones, B. C. & Perrett, D. I. (2013). Shifts in Women's Mate Preferences Across the Ovulatory Cycle: A Critique of Harris (2011) and Harris (2012). *Sex Roles*, *69*(9-10), 516–524. doi: 10.1007/s11199-013-0273-4

Gildersleeve, K., Haselton, M. G. & Fales, M. R. (2014a). Do women's mate preferences change across the ovulatory cycle? A meta-analytic review. *Psychological Bulletin*, *140*(2), 1205–1259. doi: 10.1037/a0035438

Gildersleeve, K., Haselton, M. G. & Fales, M. R. (2014b). Meta-analyses and p-curves support robust cycle shifts in women's mate preferences: Reply to Wood and Carden (2014) and Harris, Pashler, and Mickes (2014). *Psychological Bulletin*, *140*(5), 1272–1280. doi: 10.1037/a0037714

Gillen, B. (1981). Physical Attractiveness. *Personality and Social Psychology Bulletin*, *7*(2), 277–281. doi: 10.1177/014616728172015

Goodpaster, B. H., Carlson, C. L., Visser, M., Kelley, D. E., Scherzinger, A., Harris, T. B., … Newman, A. B. (2001). Attenuation of skeletal muscle and strength in the elderly: The Health ABC Study. *Journal of Applied Physiology*, *90*(6), 2157–2165.

Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika*, *40*(1), 33–51. doi: 10.1007/BF02291478

Graham, J. R., Harvey, C. R. & Puri, M. (2010, apr). *A Corporate Beauty Contest* (Working

Paper No. 15906). National Bureau of Economic Research. doi: 10.3386/w15906

Grammer, K., Fink, B. & Neave, N. (2005). Human pheromones and sexual attraction. *European Journal of Obstetrics Gynecology and Reproductive Biology*, *118*(2), 135–142. doi: 10.1016/j.ejogrb.2004.08.010

Grammer, K., Keki, V., Striebel, B., Atzmüller, M. & Fink, B. (2003). Bodies in motion: a window to the soul. In E. Voland & K. Grammer (Eds.), *Evolutionary aesthetics* (pp. 295–324). Heidelberg, Berlin, New York: Springer.

Grammer, K. & Thornhill, R. (1994). Human (Homo sapiens) facial attractiveness and sexual selection: the role of symmetry and averageness. *Journal of Comparative Psychology*, *108*(3), 233–242. doi: 10.1037/0735-7036.108.3.233

Gray, P. B., Kahlenberg, S. M., Barrett, E. S., Lipson, S. F. & Ellison, P. T. (2002). Marriage and fatherhood are associated with lower testosterone in males. *Evolution and Human Behavior*, *23*(3), 193–201.

Hamermesh, D. S. & Biddle, J. E. (1994). Beauty and the labor market. *The American Economic Review*, *84*(5), 1174–1194. doi: 10.2307/2117767

Hamilton, W. D. & Zuk, M. (1982). Heritable true fitness and bright birds - A role for parasites. *Science*, *218*(4570), 384–387.

Hancock, P. J. B. (2000). Evolving faces from principal components. *Behavior Research Methods, Instruments, & Computers*, *32*(2), 327–333. doi: 10.3758/BF03207802

Hancock, P. J. B., Burton, A. M. & Bruce, V. (1996). Face processing: Human perception and principal components analysis. *Memory & Cognition*, *24*(1), 26–40. doi: 10.3758/BF03197270

Harris, C. R. (2011). Menstrual cycle and facial preferences reconsidered. *Sex Roles*, *64*(9), 669–681. doi: 10.1007/s11199-010-9772-8

Harris, C. R. (2013). Shifts in masculinity preferences across the menstrual cycle: still not there. *Sex Roles*, *69*(9-10), 507–515. doi: 10.1007/s11199-012-0229-0

Harris, C. R., Chabot, A. & Mickes, L. (2013). Shifts in methodology and theory in menstrual cycle research on attraction. *Sex Roles*, *69*(9-10), 525–535. doi: 10.1007/s11199-013-0302-3

Harris, C. R., Pashler, H. & Mickes, L. (2014). Elastic analysis procedures: an incurable (but preventable) problem in the fertility effect literature. Comment on Gildersleeve, Haselton, and Fales (2014). *Psychological Bulletin*, *140*(5), 1260–1264. doi: 10.1037/a0036478

Hassin, R. & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology*, *78*(5), 837–852.

Havlicek, J., Cobey, K. D., Barrett, L., Klapilova, K. & Roberts, S. C. (2015). The spandrels of Santa Barbara? A new perspective on the peri-ovulation paradigm. *Behavioral Ecology*, *00*, 1–12. doi: 10.1093/beheco/arv064

Hayes, A. F. (2012). *PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper].* http://www.afhayes.com/public/process2012.pdf.

Hess, U., Adams Jr., R. B., Grammer, K. & Kleck, R. E. (2009). Face gender and emotion expression: Are angry women more like men? *Journal of Vision*, *9*(12), 1–8.

Hess, U., Adams Jr., R. B. & Kleck, R. E. (2009). The face is not an empty canvas: How facial expressions interact with facial appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3497–3504.

Holzleitner, I. J., Hunter, D. W., Tiddeman, B. P., Seck, A., Re, D. E. & Perrett, D. I. (2014). Men's facial masculinity: when (body) size matters. *Perception*, *43*(11), 1191–1202. doi: 10.1068/P7673

Innes, E. (1999). Handgrip strength testing: A review of the literature. *Australian Occupational Therapy Journal*, *46*(3), 120–140. doi: 10.1046/j.1440-1630.1999.00182.x

Jasienska, G., Ziomkiewicz, A., Ellison, P. T., Lipson, S. F. & Thune, I. (2004). Large breasts and narrow waists indicate high reproductive potential in women. *Proc. R. Soc. B*, *271*(1545), 1213–1217.

Jenkins, R., White, D., Van Montfort, X. & Burton, A. M. (2011). Variability in photos of the same face. *Cognition*, *121*(3), 313–323.

Johnston, V. S. & Franklin, M. (1993). Is beauty in the eye of the beholder? *Ethology and Sociobiology*, *14*(3), 183–199.

Johnston, V. S., Hagel, R., Franklin, M., Fink, B. & Grammer, K. (2001). Male facial attractiveness: evidence for hormone-mediated adaptive design. *Evolution and Human Behavior*, *22*(4), 251–267.

Jokela, M. (2009). Physical attractiveness and reproductive success in humans: Evidence from the late 20 century United States. *Evolution and Human Behavior*, *30*(5), 342–350. doi: 10.1016/j.evolhumbehav.2009.03.006

Jones, B. C., DeBruine, L. M., Perrett, D. I., Little, A. C., Feinberg, D. R. & Law Smith, M. J. (2008). Effects of menstrual cycle phase on face preferences. *Archives of Sexual Behavior*, *37*(1), 78–84.

Jones, B. C., Feinberg, D. R., Watkins, C. D., Fincher, C. L., Little, A. C. & DeBruine, L. M. (2013). Pathogen disgust predicts women's preferences for masculinity in men's voices, faces, and bodies. *Behavioral Ecology*, *24*(2), 373–379. doi: 10.1093/beheco/ars173

Jones, B. C., Little, A. C., Boothroyd, L. G., DeBruine, L. M., Feinberg, D. R., Law Smith, M. J., … Perrett, D. I. (2005). Commitment to relationships and preferences for femininity and apparent health in faces are strongest on days of the menstrual cycle when progesterone level is high. *Hormones and Behavior*, *48*(3), 283–290.

Jones, B. C., Little, A. C., Burt, D. M. & Perrett, D. I. (2004). When facial attractiveness is only skin deep. *Perception*, *33*(5), 569–576.

Jones, B. C., Little, A. C. & Perrett, D. I. (2003). Why are symmetrical faces more attractive? In S. P. Shohov (Ed.), *Advances in psychology research* (pp. 145–166). Nova Science Publishers.

Jones, D. (1996). An evolutionary perspective on physical attractiveness. *Evolutionary Anthropology*, *5*(3), 97–109.

Jones, D., Brace, C. L., Jankowiak, W., Laland, K. N., Musselman, L. E., Langlois, J. H., … Symons, D. (1995). Sexual selection, physical attractiveness, and facial neoteny: cross-cultural evidence and implications [and comments and reply]. *Current Anthropology*, *36*(5), 723–748.

Jones, D. & Hill, K. (1993). Criteria of facial attractiveness in five populations. *Human Nature*, *4*(3), 271–296.

Jürges, H., Avendano, M. & MacKenbach, J. P. (2008). Are different measures of self-rated health comparable? An assessment in five European countries. *European Journal of Epidemiology*, *23*(12), 773–781. doi: 10.1007/s10654-008-9287-6

Kandrik, M. & DeBruine, L. M. (2012). Self-rated attractiveness predicts preferences for opposite-sex faces, while self-rated sex-typicality predicts preferences for same-sex faces. *Journal of Evolutionary Psychology*, *10*(4), 177–186. doi: 10.1556/JEP.10.2012.4.2

Kanwisher, N. & Barton, J. J. S. (2011). The functional architecture of the face system: integrating evidence from fMRI and patient studies. In A. J. Calder, G. Rhodes, M. H. Johnson & J. V. Haxby (Eds.), *Oxford handbook of face perception* (pp. 111–129). Oxford, UK: Oxford University Press. doi: 10.1093/oxfordhb/9780199559053.013.0007

Kinsey, A. C., Pomeroy, W. B. & Martin, C. E. (1948). *Sexual behavior in the human male*. Philadelphia: W.B. Saunders.

Kleisner, K., Priplatova, L., Frost, P. & Flegr, J. (2013). Trustworthy-looking face meets brown eyes. *PLoS ONE*, *8*(1), 1–7. doi: 10.1371/journal.pone.0053285

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

Koehler, N., Simmons, L. W., Rhodes, G. & Peters, M. (2004). The relationship between sexual dimorphism in human faces and fluctuating asymmetry. *Proc. R. Soc. B*, *271*(SUPPL. 4), S233–S236.

Kokko, H., Brooks, R., Jennions, M. D. & Morley, J. (2003). The evolution of mate choice and mating biases. *Proc. R. Soc. B*, *270*(1515), 653–664. doi: 10.1098/rspb.2002.2235

Komori, M., Kawamura, S. & Ishihara, S. (2011). Multiple Mechanisms in the Perception of Face Gender: Effect of Sex-Irrelevant Features. *Journal of Experimental Psychology-Human Perception and Performance*, *37*(3), 626–633. doi: 10.1037/A0020369

Korthase, K. M. & Trenholme, I. (1982). Perceived age and perceived physical attractiveness.

Perceptual and Motor Skills, *54*, 1251–1258.

Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. (2015). *lmerTest: Tests in Linear Mixed Effects Models.* Retrieved from `http://cran.r-project.org/package=lmerTest`

Lafortuna, C. L., Maffiuletti, N. A., Agosti, F. & Sartorio, A. (2005). Gender variations of body composition, muscle strength and power output in morbid obesity. *International Journal of Obesity*, *29*(7), 833–841. doi: 10.1038/sj.ijo.0802955

Landolt, M. A., Lalumière, M. L. & Quinsey, V. L. (1995). Sex differences in intra-sex variations in human mating tactics: An evolutionary approach. *Ethology and Sociobiology*, *16*(1), 3–23.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M. & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*(3), 390–423.

Langlois, J. H. & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, *1*(2), 115–121. doi: 10.1111/j.1467-9280.1990.tb00079.x

Langlois, J. H., Roggman, L. A., Casey, R. J., Ritter, J. M., Rieser-Danner, L. A. & Jenkins, V. Y. (1987). Infant Preferences for Attractive Faces: Rudiments of a Stereotype? *Developmental Psychology*, *23*, 349–363.

Law Smith, M. J., Deady, D. K., Moore, F. R., Jones, B. C., Cornwell, R. E., Stirrat, M., … Perrett, D. I. (2012). Maternal tendencies in women are associated with estrogen levels and facial femininity. *Hormones and Behavior*, *61*(1), 12–16.

Law Smith, M. J., Perrett, D. I., Jones, B. C., Cornwell, R. E., Moore, F. R., Feinberg, D. R., … Hillier, S. G. (2006). Facial appearance is a cue to oestrogen levels in women. *Proc. R. Soc. B*, *273*(1583), 135–140. doi: 10.1098/rspb.2005.3296

Lee, A. J. & Zietsch, B. P. (2011). Experimental evidence that women's mate preferences are directly influenced by cues of pathogen prevalence and resource scarcity. *Biology Letters*, *7*(6), 892–895. doi: 10.1098/rsbl.2011.0454

Lee, A. J. & Zietsch, B. P. (2015). Women's pathogen disgust predicting preference for facial masculinity may be specific to age and study design. *Evolution and Human Behavior*, *36*, 249–255. doi: 10.1016/j.evolhumbehav.2014.12.001

Lefevre, C. E., Lewis, G. J., Bates, T. C., Dzhelyova, M., Coetzee, V., Deary, I. J. & Perrett, D. I. (2012). No evidence for sexual dimorphism of facial width-to-height ratio in four large adult samples. *Evolution and Human Behavior*, *33*(6), 623–627. doi: 10.1016/j.evolhumbehav.2012.03.002

Lewis, G. J., Lefevre, C. E. & Bates, T. C. (2012). Facial width-to-height ratio predicts achievement drive in US presidents. *Personality and Individual Differences*, *52*(7), 855–857.

Lieberman, D. E. (2011). *The Evolution of the Human Head*. Cambridge, Massachusetts; London, England: The Belknap Press of Harvard University Press.

Little, A. C., Apicella, C. L. & Marlowe, F. W. (2007). Preferences for symmetry in human faces in two cultures: Data from the UK and the Hadza, an isolated group of hunter-gatherers. *Proc. R. Soc. B*, *274*(1629), 3113–3117.

Little, A. C., Burriss, R. P., Petrie, M., Jones, B. C. & Roberts, S. C. (2013). Oral contraceptive use in women changes preferences for male facial masculinity and is associated with partner facial masculinity. *Psychoneuroendocrinology*, *38*(9), 1777–1785. doi: 10.1016/j.psyneuen.2013.02.014

Little, A. C., Burt, D. M., Penton-Voak, I. S. & Perrett, D. I. (2001). Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry in male faces. *Proc. R. Soc. B*, *268*(1462), 39–44.

Little, A. C., Burt, D. M. & Perrett, D. I. (2006). What is good is beautiful: face preference reflects desired personality. *Personality and Individual Differences*, *41*(6), 1107–1118.

Little, A. C., Cohen, D. L., Jones, B. C. & Belsky, J. (2007). Human preferences for facial masculinity change with relationship type and environmental harshness. *Behavioral Ecology and Sociobiology*, *61*(6), 967–973.

Little, A. C., Connely, J., Feinberg, D. R., Jones, B. C. & Roberts, S. C. (2011). Human preference for masculinity differs according to context in faces, bodies, voices, and smell. *Behavioral Ecology*, *22*(4), 862–868.

Little, A. C., DeBruine, L. M. & Jones, B. C. (2011). Exposure to visual cues of pathogen contagion changes preferences for masculinity and symmetry in opposite-sex faces. *Proc. R. Soc. B*, *278*(1714), 2032–2039.

Little, A. C., DeBruine, L. M. & Jones, B. C. (2013). Environment contingent preferences: Exposure to visual cues of direct male-male competition and wealth increase women's preferences for masculinity in male faces. *Evolution and Human Behavior*, *34*(3), 193–200. doi: 10.1016/j.evolhumbehav.2012.11.008

Little, A. C. & Jones, B. C. (2012). Variation in facial masculinity and symmetry preferences across the menstrual cycle is moderated by relationship context. *Psychoneuroendocrinology*, *37*(7), 999–1008.

Little, A. C., Jones, B. C., DeBruine, L. M. & Feinberg, D. R. (2008). Symmetry and sexual dimorphism in human faces: Interrelated preferences suggest both signal quality. *Behavioral Ecology*, *19*(4), 902–908.

Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M. & Perrett, D. I. (2002). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proc. R. Soc. B*, *269*(1496), 1095–1100.

Little, A. C., Jones, B. C., Waitt, C., Tiddeman, B. P., Feinberg, D. R., Perrett, D. I., … Marlowe,

F. W. (2008). Symmetry is related to sexual dimorphism in faces: Data across culture and species. *PLoS ONE*, *3*(5).

Little, A. C. & Mannion, H. (2006). Viewing attractive or unattractive same-sex individuals changes self-rated attractiveness and face preferences in women. *Animal Behaviour*, *72*(5), 981–987.

Little, A. C., Saxton, T. K., Roberts, S. C., Jones, B. C., DeBruine, L. M., Vukovic, J., ... Chenore, T. (2010). Women's preferences for masculinity in male faces are highest during reproductive age range and lower around puberty and post-menopause. *Psychoneuroendocrinology*, *35*(6), 912–920.

Low, B. S. (1990). Marriage systems and pathogen stress in human societies. *Integrative and Comparative Biology*, *30*(2), 325–340.

Lüdecke, D. (2015). *sjPlot: Data Visualization for Statistics in Social Science.* Retrieved from `http://cran.r-project.org/package=sjPlot`

Mazur, A. & Booth, A. (1998). Testosterone and dominance in men. *Behavioral and Brain Sciences*, *21*(03), 353–363. doi: 10.1017/S0140525X98001228

McCormick, C. M. (2013). Watch where and how you stick pins when playing with voodoo correlations. *Journal of General Psychology*, *140*(1), 82–86. doi: 10.1080/00221309.2012.737872

McEwen, B. S. (1988). Actions of sex hormones on the brain: 'organization' and 'activation' in relation to functional teratology. In G. J. Boer, M. G. P. Feenstra, M. Mirmiran, D. F. Swaab & F. Van Haaren (Eds.), *Progress in brain research* (Vol. 73, pp. 121–134). Elsevier. doi: 10.1016/S0079-6123(08)60501-4

Mehta, P. H. & Josephs, R. A. (2006). Testosterone change after losing predicts the decision to compete again. *Hormones and Behavior*, *50*(5), 684–692.

Mignault, A. & Chaudhuri, A. (2003). The many faces of a neutral face: head tilt and perception of dominance and emotion. *Journal of Nonverbal Behavior*, *27*(2), 111–132.

Miller, G. F. (2001). *The mating mind: how sexual selection shaped the evolution of human nature*. Anchor Books.

Mitteroecker, P., Gunz, P., Windhager, S. & Schaefer, K. (2013). A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. *Hystrix, the Italian Journal of Mammalogy*, *24*(1).

Mitteroecker, P., Windhager, S., Müller, G. B. & Schaefer, K. (2015). The morphometrics of "masculinity" in human faces. *PLOS ONE*, *10*(2), e0118374. doi: 10.1371/journal.pone.0118374

Moller, A. P. & Thornhill, R. (1998). Bilateral symmetry and sexual selection: A meta-analysis. *American Naturalist*, *151*(2), 174–192.

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.

Moore, F. R., Al Dujaili, E. A. S., Cornwell, R. E., Law Smith, M. J., Lawson, J. F., Sharp, M. & Perrett, D. I. (2011). Cues to sex- and stress-hormones in the human male face: Functions of glucocorticoids in the immunocompetence handicap hypothesis. *Hormones and Behavior*, *60*(3), 269–274.

Moore, F. R., Cassidy, C., Law Smith, M. J. & Perrett, D. I. (2006). The effects of female control of resources on sex-differentiated mate preferences. *Evolution and Human Behavior*, *27*(3), 193–205. doi: 10.1016/j.evolhumbehav.2005.08.003

Moore, F. R., Cornwell, R. E., Law Smith, M. J., Al Dujaili, E. A. S., Sharp, M. & Perrett, D. I. (2011). Evidence for the stress-linked immunocompetence handicap hypothesis in human male faces. *Proc. R. Soc. B*, *278*(1706), 774–780.

Muehlenbein, M. P. & Bribiescas, R. G. (2005). Testosterone-mediated immune functions and male life histories. *American Journal of Human Biology*, *17*(5), 527–558.

Mueller, U. & Mazur, A. (1997). Facial dominance in Homo sapiens as honest signaling of male quality. *Behavioral Ecology*, *8*(5), 569–579.

Munoz-Reyes, J. A., Gil-Burmann, C., Fink, B. & Turiegano, E. (2012). Physical strength, fighting ability, and aggressiveness in adolescents. *American Journal of Human Biology*, *24*(5), 611–617. doi: 10.1002/Ajhb.22281

Neave, N., Laing, S., Fink, B. & Manning, J. T. (2003). Second to fourth digit ratio, testosterone and perceived male dominance. *Proc. R. Soc. B*, *270*(1529), 2167–2172. doi: 10.1098/rspb.2003.2502

Neth, D. & Martinez, A. M. (2009). Emotion perception in emotionless face images suggests a norm-based representation. *Journal of Vision*, *9*(1), 5.1–11. doi: 10.1167/9.1.5

Newman, A. B., Haggerty, C. L., Goodpaster, B., Harris, T., Kritchevsky, S., Nevitt, M., … Visser, M. (2003). Strength and muscle quality in a well-functioning cohort of older adults: The Health, Aging and Body Composition Study. *Journal of the American Geriatrics Society*, *51*(3), 323–330. doi: 10.1046/j.1532-5415.2003.51105.x

O'Connor, J. J. M., Feinberg, D. R., Fraccaro, P. J., Borak, D. J., Tigue, C. C., Re, D. E., … Tiddeman, B. P. (2012). Female preferences for male vocal and facial masculinity in videos. *Ethology*, *118*(4), 321–330.

Olivola, C. Y. & Todorov, A. (2010a). *Elected in 100 milliseconds: Appearance-based trait inferences and voting* (Vol. 34) (No. 2). doi: 10.1007/s10919-009-0082-1

Olivola, C. Y. & Todorov, A. (2010b). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, *46*(2), 315–324. doi: 10.1016/j.jesp.2009.12.002

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), 11087–11092.

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

Oppenheimer, D. M., Meyvis, T. & Davidenko, N. (2009). Instructional manipulation checks: detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872. doi: 10.1016/j.jesp.2009.03.009

O'Toole, A. J. (2011). Cognitive and computational approaches to face recognition. In A. J. Calder, G. Rhodes, M. H. Johnson & J. V. Haxby (Eds.), *The oxford handbook of face perception* (pp. 16–30). Oxford, UK: Oxford University Press.

O'Toole, A. J., Abdi, H., Deffenbacher, K. A. & Valentin, D. (1993). Low-dimensional representation of faces in higher dimensions of the face space. *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, *10*(3), 405–411.

Pawlowski, B. & Dunbar, R. I. M. (1999). Impact of market value on human mate choice decisions. *Proc. R. Soc. B*, *266*(1416), 281–285.

Penton-Voak, I. S., Jacobson, A. & Trivers, R. (2004). Populational differences in attractiveness judgements of male and female faces: Comparing British and Jamaican samples. *Evolution and Human Behavior*, *25*(6), 355–370.

Penton-Voak, I. S., Jones, B. C., Little, A. C., Baker, S., Tiddeman, B. P., Burt, D. M. & Perrett, D. I. (2001). Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. *Proc. R. Soc. B*, *268*(1476), 1617–1623.

Penton-Voak, I. S., Little, A. C., Jones, B. C., Burt, D. M., Tiddeman, B. P. & Perrett, D. I. (2003). Female condition influences preferences for sexual dimorphism in faces of male humans (Homo sapiens). *Journal of Comparative Psychology*, *117*(3), 264–271.

Penton-Voak, I. S. & Perrett, D. I. (2000). Female preference for male faces changes cyclically: Further evidence. *Evolution and Human Behavior*, *21*(1), 39–48.

Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K. & Minamisawa, R. (1999). Menstrual cycle alters face preference. *Nature*, *399*(6738), 741–742.

Penton-Voak, I. S., Pound, N., Little, A. C. & Perrett, D. I. (2006). Personality judgments from natural and composite facial images: more evidence for a "kernel of truth" in social perception. *Social Cognition*, *24*(5), 607–640. doi: 10.1521/soco.2006.24.5.607

Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A. & Edwards, R. (1999). Symmetry and human facial attractiveness. *Evolution and Human Behavior*, *20*(5), 295–307. doi: 10.1016/S1090-5138(99)00014-8

Perrett, D. I., Lee, K. J., Penton-Voak, I. S., Rowland, D., Yoshikawa, S., Burt, D. M., … Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, *394*(6696), 884–887.

Perrett, D. I., May, K. A. & Yoshikawa, S. (1994). Facial shape and judgements of female attractiveness. *Nature*, *368*(6468), 239–242. doi: 10.1038/368239a0

Perrett, D. I. & Penton-Voak, I. S. (1999). Reply to Meyer & Quong 1999. *Nature*, *397*(6721), 661–662.

Pflüger, L. S., Oberzaucher, E., Katina, S., Holzleitner, I. J. & Grammer, K. (2012). Cues to fertility: Perceived attractiveness and facial shape predict reproductive success. *Evolution and Human Behavior*, *33*(6), 708–714. doi: 10.1016/j.evolhumbehav.2012.05.005

Pollet, T. V., der Meij, L. V., Cobey, K. D. & Buunk, A. P. (2011). Testosterone levels and their associations with lifetime number of opposite sex partners and remarriage in a large sample of American elderly men and women. *Hormones and Behavior*, *60*(1), 72–77.

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. doi: 10.1037/0008-400X.40.3.171

Pound, N., Penton-Voak, I. S. & Surridge, A. K. (2009). Testosterone responses to competition in men are related to facial masculinity. *Proc. R. Soc. B*, *276*(1654), 153–159.

Puts, D. A. (2010). Beauty and the beast: Mechanisms of sexual selection in humans. *Evolution and Human Behavior*, *31*(3), 157–175.

Puts, D. A., Jones, B. C. & DeBruine, L. M. (2012). Sexual selection on human faces and voices. *Journal of Sex Research*, *49*(2-3), 227–243.

R Development Core Team. (2015). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from `http://www.r-project.org`

Rantala, M. J., Coetzee, V., Moore, F. R., Skrinda, I., Kecko, S., Krama, T., … Krams, I. (2013). Adiposity, compared with masculinity, serves as a more valid cue to immunocompetence in human mate choice. *Proc. R. Soc. B*, *280*(1751), 20122495. doi: 10.1098/rspb.2012.2495

Rantala, M. J., Moore, F. R., Skrinda, I., Krama, T., Kivleniece, I., Kecko, S. & Krams, I. (2012). Evidence for the stress-linked immunocompetence handicap hypothesis in humans. *Nature Communications*, *3*.

Re, D. E., Coetzee, V., Xiao, D., Buls, D., Tiddeman, B. P., Boothroyd, L. G. & Perrett, D. I. (2011). Viewing heavy bodies enhances preferences for facial adiposity. *Journal of Evolutionary Psychology*, *9*(4), 295–308.

Re, D. E., Hunter, D. W., Coetzee, V., Tiddeman, B. P., Xiao, D., DeBruine, L. M., … Perrett, D. I. (2013). Looking like a leader - facial shape predicts perceived height and leadership ability. *PLoS ONE*, *8*(12), e80957. doi: 10.1371/journal.pone.0080957

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). doi: 10.1371/journal.pone.0034293

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology*, *57*, 199–226. doi: 10.1146/annurev.psych.57.102904.190208

Rhodes, G., Chan, J., Zebrowitz, L. A. & Simmons, L. W. (2003). Does sexual dimorphism in human faces signal health? *Proc. R. Soc. B*, *270*(SUPPL. 1), S93–S95.

Rhodes, G., Hickford, C. & Jeffery, L. (2000). Sex-typicality and attractiveness: are supermale and superfemale faces super-attractive? *The British journal of psychology*, *91*, Pt 1/.

Rhodes, G., Simmons, L. W. & Peters, M. (2005). Attractiveness and sexual behavior: Does attractiveness enhance mating success? *Evolution and Human Behavior*, *26*(2), 186–201.

Rhodes, G. & Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychological Science*, *7*(2), 105–110.

Rhodes, G., Yoshikawa, S., Palermo, R., Simmons, L. W., Peters, M., Lee, K., … Crawford, J. R. (2007). Perceived health contributes to the attractiveness of facial symmetry, averageness, and sexual dimorphism. *Perception*, *36*(8), 1244–1252.

Rikowski, A. & Grammer, K. (1999). Human body odour, symmetry and attractiveness. *Proc. R. Soc. B*, *266*(1422), 869–874.

Ritts, V., Patterson, M. L. & Tubbs, M. E. (1992). Expectations, impressions, and judgments of physically attractive students: a review. *Review of Educational Research*, *62*(4), 413–426.

Roberts, M. L., Buchanan, K. L. & Evans, M. R. (2004). Testing the immunocompetence handicap hypothesis: A review of the evidence. *Animal Behaviour*, *68*(2), 227–239.

Roberts, S. C., Little, A. C., Gosling, L. M., Perrett, D. I., Carter, V., Jones, B. C., … Petrie, M. (2005). MHC-heterozygosity and human facial attractiveness. *Evolution and Human Behavior*, *26*(3), 213–226.

Rowe, R., Maughan, B., Worthman, C. M., Costello, E. J. & Angold, A. (2004). Testosterone, antisocial behavior, and social dominance in boys: Pubertal development and biosocial interaction. *Biological Psychiatry*, *55*(5), 546–552.

Rowland, D. A. & Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications*, *15*(5), 70–76. doi: 10.1109/38.403830

Rudoy, J. D. & Paller, K. a. (2009). Who can you trust? Behavioral and neural differences between perceptual and memory-based influences. *Frontiers in human neuroscience*, *3*(August), 16. doi: 10.3389/neuro.09.016.2009

Rule, N. O. & Ambady, N. (2008). The face of success. *Psychological Science*, *19*(2), 109–111. doi: 10.1111/j.1467-9280.2008.02054.x

Rule, N. O. & Ambady, N. (2009). She's got the look: Inferences from female chief executive officers' faces predict their success. *Sex Roles*, *61*(9-10), 644–652. doi: 10.1007/s11199-009-9658-9

Rule, N. O., Ambady, N., Adams Jr., R. B., Ozono, H., Nakashima, S., Yoshikawa, S. & Watabe, M. (2010). Polling the face: Prediction and consensus across cultures. *Journal of Personality and Social Psychology*, *98*(1), 1–15. doi: 10.1037/a0017673

Sacco, D. F., Jones, B. C., DeBruine, L. M. & Hugenberg, K. (2012). The roles of sociosexual orientation and relationship status in women's face preferences. *Personality and Individual Differences*, *53*(8), 1044–1047. doi: 10.1016/j.paid.2012.07.023

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

Said, C. P. & Todorov, A. (2011). A statistical model of facial attractiveness. *Psychological Science*, *22*(9), 1183–1190.

Sartorio, A., Agosti, F., De Col, A. & Lafortuna, C. L. (2006). Age- and gender-related variations of leg power output and body composition in severely obese children and adolescents. *Journal of Endocrinological Investigation*, *29*(1), 48–54.

Sartorio, A., Lafortuna, C. L., Pogliaghi, S. & Trecate, L. (2002). The impact of gender, body dimension and body composition on hand-grip strength in healthy children. *Journal of Endocrinological Investigation*, *25*(5), 431–435.

Sartorio, A., Proietti, M., Marinone, P. G., Agosti, F., Adorni, F. & Lafortuna, C. L. (2004). Influence of gender, age and BMI on lower limb muscular power output in a large population of obese men and women. *International journal of obesity and related metabolic disorders*, *28*(1), 91–98. doi: 10.1038/sj.ijo.0802433

Saxton, T. K., Little, A. C., DeBruine, L. M., Jones, B. C. & Roberts, S. C. (2009). Adolescents' preferences for sexual dimorphism are influenced by relative exposure to male and female faces. *Personality and Individual Differences*, *47*(8), 864–868.

Schaal, B., Tremblay, R. E., Soussignan, R. & Susman, E. J. (1996). Male testosterone linked to high social dominance but low physical aggression in early adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(10), 1322–1330.

Schaefer, K., Mitteroecker, P., Gunz, P., Bernhard, M. & Bookstein, F. L. (2004). Craniofacial sexual dimorphism patterns and allometry among extant hominids. *Annals of Anatomy*, *186*(5-6), 471–478.

Scheib, J. E., Gangestad, S. W. & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proc. R. Soc. B*, *266*(1431), 1913–1917.

Scott, I. M. L., Clark, A. P., Boothroyd, L. G. & Penton-Voak, I. S. (2013). Do men's faces really signal heritable immunocompetence? *Behavioral Ecology*, *24*(3), 579–589. doi: 10.1093/beheco/ars092

Scott, I. M. L., Pound, N., Stephen, I. D., Clark, A. P. & Penton-Voak, I. S. (2010). Does masculinity matter? The contribution of masculine face shape to male attractiveness in humans. *PLoS ONE*, *5*(10). doi: 10.1371/journal.pone.0013585

Scott, I. M. L., Swami, V., Josephson, S. C. & Penton-Voak, I. S. (2008). Context-dependent preferences for facial dimorphism in a rural Malaysian population. *Evolution and Human Behavior*, *29*(4), 289–296.

Secord, P. F. (1958). Facial features and inference processes in interpersonal perception. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior* (pp. 300–315). Stanford, CA: Stanford University Press. doi: 10.1097/00007611-195903000-00037

Sell, A. N., Cosmides, L. & Tooby, J. (2014). The human anger face evolved to enhance cues of strength. *Evolution and Human Behavior*, *35*(5), 425–429. doi: 10.1016/j.evolhumbehav.2014.05.008

Sell, A. N., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C. & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proc. R. Soc. B*, *276*(1656), 575–584. doi: 10.1098/rspb.2008.1177

Sigall, H. & Ostrove, N. (1975). Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridic judgment. *Journal of Personality and Social Psychology*, *31*(3), 410–414.

Simmons, Z. L. & Roney, J. R. (2011). Variation in CAG repeat length of the androgen receptor gene predicts variables associated with intrasexual competitiveness in human males. *Hormones and Behavior*, *60*(3), 306–312. doi: 10.1016/j.yhbeh.2011.06.006

Sirovich, L. & Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America A*, *4*(3), 519–524. doi: 10.1364/JOSAA.4.000519

Slice, D. E. (2005). Modern Morphometrics. In D. E. Slice (Ed.), *Modern morphometrics in physical anthropology* (pp. 1–45). New York: Kluwer Academic / Plenum Publishers.

Slice, D. E. (2007). Geometric morphometrics. *Annual Review of Anthropology*, *36*, 261–281. doi: 10.1146/annurev.anthro.36.081804.120613

Smith, F. G., Jones, B. C. & DeBruine, L. M. (2010). Individual differences in empathizing and systemizing predict variation in face preferences. *Personality and Individual Differences*, *49*(6), 655–658.

Smith, F. G., Jones, B. C., Little, A. C., DeBruine, L. M., Welling, L. L. M., Vukovic, J. & Conway, C. A. (2009). Hormonal contraceptive use and perceptions of trust modulate the effect of relationship context on women's preferences for sexual dimorphism in male face shape. *Journal of Evolutionary Psychology*, *7*(3), 195–210. doi: 10.1556/JEP.7.2009.3.1

Stephen, I. D., Scott, I. M. L., Coetzee, V., Pound, N., Perrett, D. I. & Penton-Voak, I. S. (2012). Cross-cultural effects of color, but not morphological masculinity, on perceived attractiveness of men's faces. *Evolution and Human Behavior*, *33*(4), 260–267. doi: 10.1016/j.evolhumbehav.2011.10.003

Stirrat, M., Stulp, G. & Pollet, T. V. (2012). Male facial width is associated with death by contact violence: Narrow-faced males are more likely to die from contact violence. *Evolution and Human Behavior*, *33*(5), 551–556. doi: 10.1016/j.evolhumbehav.2012.02.002

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

Swaddle, J. P. & Reierson, G. W. (2002). Testosterone increases perceived dominance but not attractiveness in human males. *Proc. R. Soc. B*, *269*(1507), 2285–2289.

Swami, V. & Tovée, M. J. (2005). Male physical attractiveness in Britain and Malaysia: a cross-cultural study. *Body Image*, *2*(4), 383–393. doi: 10.1016/j.bodyim.2005.08.001

Symons, D. (1995). Beauty is in the adaptations of the beholder: The evolutionary psychology of human female sexual attractiveness. In P. R. Abramson & S. D. Pinkerton (Eds.), *Sexual nature, sexual culture* (pp. 80–118). Chicago and London: The University of Chicago Press.

Thornhill, R. & Gangestad, S. W. (1993). Human facial beauty - averageness, symmetry, and parasite resistance. *Human Nature*, *4*(3), 237–269.

Thornhill, R. & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences*, *3*(12), 452–460.

Thornhill, R. & Gangestad, S. W. (2006). Facial sexual dimorphism, developmental stability, and susceptibility to disease in men and women. *Evolution and Human Behavior*, *27*(2), 131–144.

Thornhill, R. & Moller, A. P. (1997). Developmental stability, disease and medicine. *Biological Reviews of the Cambridge Philosophical Society*, *72*(4), 497–548.

Tiddeman, B. P., Burt, D. M. & Perrett, D. I. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, *21*(5), 42–50.

Tiddeman, B. P., Duffy, N. & Rabey, G. (2000). Construction and visualisation of three-dimensional facial statistics. *Computer Methods and Programs in Biomedicine*, *63*(1), 9–20.

Todorov, A., Dotsch, R., Porter, J. M., Oosterhof, N. N. & Falvello, V. B. (2013). Validation of data-driven computational models of social perception of faces. *Emotion*, *13*(4), 724–38. doi: 10.1037/a0032335

Todorov, A., Mandisodza, A. N., Goren, A. & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623–1626. doi: 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, 1–46. doi: 10.1146/annurev-psych-113011-143831
Todorov, A. & Oosterhof, N. N. (2011). Modeling Social Perception of Faces. *Signal Processing Magazine, IEEE*, *28*(March), 117–122. doi: 10.1109/MSP.2010.940006
Todorov, A., Pakrashi, M. & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, *27*(6), 813–833. doi: 10.1521/soco.2009.27.6.813
Todorov, A. & Porter, J. M. (2014). Misleading first impressions: different for different facial images of the same person. *Psychological Science*, *25*(7), 1404–1417. doi: 10.1177/0956797614532474
Toscano, H., Schubert, T. W. & Sell, A. N. (2014). Judgments of dominance from the face track physical strength. *Evolutionary Psychology*, *12*(1), 1–18.
Trampisch, U. S., Franke, J., Jedamzik, N., Hinrichs, T. & Platen, P. (2012). Optimal Jamar Dynamometer Handle Position to Assess Maximal Isometric Hand Grip Strength in Epidemiological Studies. *The Journal of Hand Surgery*, *37*(11), 2368–2373. doi: 10.1016/j.jhsa.2012.08.014
Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the descent of men, 1871-1971* (pp. 136–179). Chicago: Aldine Publishing Company.
Turk, M. & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*(1), 71–86.
Tybur, J. M., Lieberman, D. & Griskevicius, V. (2009). Microbes, mating, and morality: individual differences in three functional domains of disgust. *Journal of Personality and Social Psychology*, *97*(1), 103–122. doi: 10.1037/a0015474
Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly journal of experimental psychology. A, Human experimental psychology*, *43*(2), 161–204. doi: 10.1080/14640749108400966
Valenzano, D. R., Mennucci, A., Tartarelli, G. & Cellerino, A. (2006). Shape analysis of female facial attractiveness. *Vision Research*, *46*(8-9), 1282–1291. doi: 10.1016/j.visres.2005.10.024
Vilaca, K. H. C., Carneiro, J. A. O., Ferriolli, E., Lima, N. K. D., de Paula, F. J. A. & Moriguti, J. C. (2014). Body composition, physical performance and muscle quality of active elderly women. *Archives of Gerontology and Geriatrics*, *59*(1), 44–48. doi: 10.1016/j.archger.2014.02.004
von Rueden, C., Gurven, M. & Kaplan, H. (2011). Why do men seek status? Fitness payoffs to dominance and prestige. *Proc. R. Soc. B*, *278*(1715), 2223–2232. doi: 10.1098/rspb.2010.2145
Vukovic, J., Feinberg, D. R., Jones, B. C., DeBruine, L. M., Welling, L. L. M., Little, a. C. & Smith, F. G. (2008). Self-rated attractiveness predicts individual differences in women's

preferences for masculine men's voices. *Personality and Individual Differences*, *45*(6), 451–456. doi: 10.1016/j.paid.2008.05.013

Walker, M., Jiang, F., Vetter, T. & Sczesny, S. (2011). Universals and cultural differences in forming personality trait judgments from faces. *Social Psychological and Personality Science*, *2*(6), 609–617. doi: 10.1177/1948550611402519

Walker, M. & Vetter, T. (2009). Portraits made to measure: manipulating social judgments about individuals with a statistical face model. *Journal of Vision*, *9*(11), 12.1–13. doi: 10.1167/9.11.12

Walster, E., Aronson, V., Abrahams, D. & Rottmann, L. (1966). Importance of physical attractiveness in dating behavior. *Journal of Personality and Social Psychology*, *4*(5), 508–516.

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. doi: 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.

Watkins, C. D., Quist, M. C., Smith, F. G., DeBruine, L. M. & Jones, B. C. (2012). Individual differences in women's perceptions of other women's dominance. *European Journal of Personality*, *26*(1), 79–86.

Waynforth, D., Delwadia, S. & Camm, M. (2005). The influence of women's mating strategies on preference for masculine facial architecture. *Evolution and Human Behavior*, *26*(5), 409–416.

Wedekind, C. & Folstad, I. (1994). Adaptive or Nonadaptive Immunosuppression by Sex Hormones? *The American Naturalist*, *143*(5), 936–938.

Welling, L. L. M., DeBruine, L. M., Little, A. C. & Jones, B. C. (2009). Extraversion predicts individual differences in women's face preferences. *Personality and Individual Differences*, *47*(8), 996–998.

Welling, L. L. M., Jones, B. C. & DeBruine, L. M. (2008). Sex drive is positively associated with women's preferences for sexual dimorphism in men's and women's faces. *Personality and Individual Differences*, *44*(1), 161–170.

Whitehead, R. D., Re, D. E., Xiao, D., Ozakinci, G. & Perrett, D. I. (2012). You are what you eat: within-subject increases in fruit and vegetable consumption confer beneficial skin-color changes. *PLoS ONE*, *7*(3), e32988. doi: 10.1371/journal.pone.0032988

WHO. (2013). *BMI Classification* (Vol. 2013) (No. July 11). Global Database on Body Mass Index. Retrieved from `http://apps.who.int/bmi/index.jsp?introPage=intro{_}3.html`

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

Williams, G. C. (1966). *Adaptation and natural selection*. Princeton, NJ: Princeton University Press.

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. doi: 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. doi: 10.1177/0956797615590992

Wind, A., Takken, T., Helders, P. & Engelbert, R. (2010). Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? *European Journal of Pediatrics*, *169*(3), 281–287. doi: 10.1007/s00431-009-1010-4

Windhager, S., Patocka, K. & Schaefer, K. (2013). Body fat and facial shape are correlated in female adolescents. *American Journal of Human Biology*. doi: 10.1002/ajhb.22444

Windhager, S., Schaefer, K. & Fink, B. (2011). Geometric morphometrics of male facial shape in relation to physical strength and perceived attractiveness, dominance, and masculinity. *American Journal of Human Biology*, *23*(6), 805–814.

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

Wood, W. (2014). Author reply: Once again, menstrual cycles and mate preferences. *Emotion Review*, *6*(3), 258–260. doi: 10.1177/1754073914523053

Wood, W. & Carden, L. (2014). Elusiveness of menstrual cycle effects on mate preferences: comment on Gildersleeve, Haselton, and Fales (2014). *Psychological Bulletin*, *140*(5), 1265–1271. doi: 10.1037/a0036722This

Wood, W., Kressel, L., Joshi, P. D. & Louie, B. (2014). Meta-analysis of menstrual cycle effects on women's mate preferences. *Emotion Review*, *6*(3), 229–249. doi: 10.1177/1754073914523073

Zahavi, A. (1975). Mate selection: a selection for a handicap. *Journal of Theoretical Biology*, *53*(1), 205–214.

Zahavi, A. & Zahavi, A. (1997). *The Handicap Principle: A Missing Piece of Darwin's Puzzle* (Vol. 117) (No. 1). New York, Oxford: Oxford University Press.

Zebrowitz, L. A. (1996). Physical appearance as a basis of stereotyping. In C. N. Macrae, C. Stangor & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 79–120). New York: The Guilford Press.

Zebrowitz, L. A. (2011). Ecological and Social Approaches to Face Perception. In A. J. Calder, G. Rhodes, M. H. Johnson & J. V. Haxby (Eds.), *The oxford handbook of face perception* (pp. 31–50). Oxford, UK: Oxford University Press. doi: 10.1093/oxfordhb/9780199559053.013.0003

Zebrowitz, L. A. & Collins, M. A. (1997). Accurate social perception at zero acquaintance: the affordances of a Gibsonian approach. *Personality and Social Psychology Review*, *1*(3), 204–223. doi: http://dx.doi.org/10.1207/s15327957pspr0103{_}2

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. doi: 10.1207/S15327957pspr0703{_}01

Zebrowitz, L. A., Hall, J. A., Murphy, N. A. & Rhodes, G. (2002). Looking smart and looking good: facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin*, *28*(2), 238–249. doi: 10.1177/0146167202282009

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. doi: 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. doi: 10.1037/a0017990

Zebrowitz, L. A. & McDonald, S. M. (1991). *The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts* (Vol. 15) (No. 6). doi: 10.1007/BF01065855

Zebrowitz, L. A. & Montepare, J. M. (2005). Appearance DOES matter. *Science*, *308*(5728), 1565–1566. doi: 10.1126/science.1114170

Zebrowitz, L. A. & Montepare, J. M. (2008). Social psychological face perception: why appearance matters. *Social and Personality Psychology Compass*, *2*(3), 1497. doi: 10.1111/j.1751-9004.2008.00109.x

Zebrowitz, L. A. & Rhodes, G. (2004). Sensitivity to "bad genes" and the anomalous face overgeneralization effect: cue validity, cue utilization, and accuracy in judging intelligence and health. *Journal of Nonverbal Behavior*, *28*(3), 167–185.

Zebrowitz, L. A., Tenenbaum, D. R. & Goldstein, L. H. (1991). The impact of job applicants' facial maturity, gender, and academic achievenment on hiring recommendations. *Journal of Applied Social Psychology*, *21*(7), 525–548.

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

Zebrowitz McArthur, L. & Berry, D. S. (1987). Cross-cultural agreement in perceptinos of babyfaced adults. *Journal of Cross-Cultural Psychology*, *18*(2), 165–192.

Zhang, P., Peterson, M., Su, G. L. & Wang, S. C. (2015). Visceral adiposity is negatively associated with bone density and muscle attenuation. *American Journal of Clinical Nutrition*, *101*(2), 337–343. doi: 10.3945/ajcn.113.081778

Zietsch, B. P., Lee, A. J., Sherlock, J. M. & Jern, P. (2015). Variation in women's preferences regarding male facial masculinity is better explained by genetic differences than by previously identified context-dependent effects. *Psychological Science*. doi: 10.1177/0956797615591770

Zilioli, S., Sell, A. N., Stirrat, M., Jagore, J., Vickerman, W. & Watson, N. V. (2014). Face of a fighter: Bizygomatic width as a cue of formidability. *Aggressive Behavior*. doi: 10.1002/ab.21544

Appendix

A. Face Set 2 questionnaire data

1. Please indicate your age (drop-down selection)
2. Please select your gender (drop-down selection)
3. Please indicate your ethnic group (drop-down selection)
4. Please indicate your age at your first period/first shave (drop-down selection)
5. Compared to my peers at the time, my physical growth and development was: 1 Much earlier - 7 Much later
6. Compared to my peers at the time, I went through puberty: 1 Much slower - 7 Much faster
7. At what age did you gain the most height? (drop-down selection)
8. At what age did you reach your final height? (drop-down selection)
9. Compared to other people, how healthy do you think you are in general? 1 Not at all healthy - 7 Extremely healthy
10. How many times and days in the past 3 years have you had a respiratory infection (cold, flu) or stomach or intestinal flu?
    Respiratory infections: Number ___________ Days infected ___________
    Stomach and intestinal infections: Number ___________ Days infected ___________
11. For how many infections in the past 3 years did you take antibiotics?
    Number of times antibiotics used: ___________
12. Please use the following 7-point scale to indicate how much you agree with the following statement:
    If I were in a physical fight with a person of the same sex, I would probably win. (1 Strongly disagree - 7 Strongly agree)
13. I am physically stronger than ___________% of others of my sex.

14. A socially dominant person tells other people what to do, is respected, influential, and often a leader, while submissive people are not influential or assertive and are usually directed by others. Based on the above statement, do you consider yourself to be: 1 Highly submissive - 7 Highly dominant

15. Please indicate how accurately each statement describes you (1 Very inaccurate - 7 Very accurate):

    a) I try to surpass others' accomplishments
    b) I try to outdo others
    c) I am quick to correct others
    d) I impose my will on others
    e) I demand explanations from others
    f) I want to control the conversation
    g) I am not afraid of providing criticism
    h) I challenge others' point of view
    i) I lay down the law to others
    j) I put people under pressure
    k) I hate to seem pushy

16. Please indicate how much you agree with each of these statements (1 Strongly disagree - 7 Strongly agree):

    a) Members of the opposite sex that I like, tend to like me back.
    b) Members of the opposite sex notice me.
    c) I receive many compliments from members of the opposite sex.
    d) Members of the opposite sex are not very attracted to me.
    e) I receive sexual invitations from members of the opposite sex.
    f) Members of the opposite sex are attracted to me.
    g) I can have as many sexual partners as I choose.
    h) I do not receive many compliments from members of the opposite sex.

17. I've answered all questions truthfully. (Yes/No)

B. Delineation template

Table B.1.: List of definitions and operationalization of used landmarks (following Farkas, 1994). Italics indicate names of traditional anthropometric landmarks.

Landmark	Definition
1	*Nasion*; on midsagittal plane, in lateral view on lowest point above the nose
2	Centre of right pupil
3	Centre of left pupil
4	*Exocanthion* right; outer corner of the right eye fissure where eyelids meet
5	*Endocanthion* right; inner corner of the eye fissure where eyelids meet
6	Highest point of right iris
7	Lowest point of right iris
8	*Endocanthion* left
9	*Exocanthion* left
10	Highest point of left iris
11	Lowest point of left iris
12	*Alare* right; most lateral point on the right ala
13	*Alare* left
14	*Cheilion* right; right corner of the mouth where the outer edges of upper and lower vermillion meet
15	*Cheilion* left
16	*Labrale superius*; midpoint of the upper vermillion line
17	*Labrale inferius*; midpoint of the lower vermillion line
18	Mid-cleft of upper vermillion
19	Mid-cleft of lower vermillion
20	*Trichion*, midpoint of the hairline
21	*Gnathion*; midpoint of chin
22	Frontal view: right outermost feature of face along the horizontal axis of the mouth. Lateral view: turning point of Ramus and Corpus mandibulae

(continued)

Landmark	Definition
23	Frontal view: left outermost feature of face along the horizontal axis of the mouth. Lateral view: turning point of Ramus and Corpus mandibulae
24	*Glabella*; on midsagittal plane, joins the superciliary ridges; lateral view: most protuberant point
25	Tip of the nose; lateral view: most protuberant point on nose
26	*Subnasale*; on the local midline of the junction formed by lower border of nasal septum and cutaneous portion of upper lip
27	Lateral view: deepest point between lip red and chin
28	Lateral view: most protuberant point of chin
29	Lowest point of attachment of right external ear to the face
30	Lowest point of attachment of left external ear to the face
31	*Superciliare mediale* right; most medial point of eyebrow
32	Midpoint of right eyebrow (horizontally and vertically)
33	*Supercilare laterale* right; most lateral point of right eyebrow
34	*Superciliare mediale* left
35	Midpoint of left eyebrow
36	*Superciliare laterale* left
37	*Crista philtrum* right; right crest of the philtrum, i.e. the vertical groove in the median portion of upper lip, located on the vermillion border
38	*Crista philtrum* left
39, 40	Evenly spaced between 21 and 22 along jaw line
41, 42	Evenly spaced between 21 and 23 along jaw line
43	On midsaggital plane beneath chin
44	Lateral view: right intersection of sternocleidomastoid muscle and jaw (excluded for shape analyses in Study 3)
45	Lateral view: left intersection of sternocleidomastoid muscle and jaw (excluded for shape analyses in Study 3)
46	Right intersection of pupil line and hairline
47	Left intersection of pupil line and hairline
48, 49	Evenly spaced along hairline between 20 and 46
50, 51	Evenly spaced along hairline between 20 and 47

C. Principal components of shape

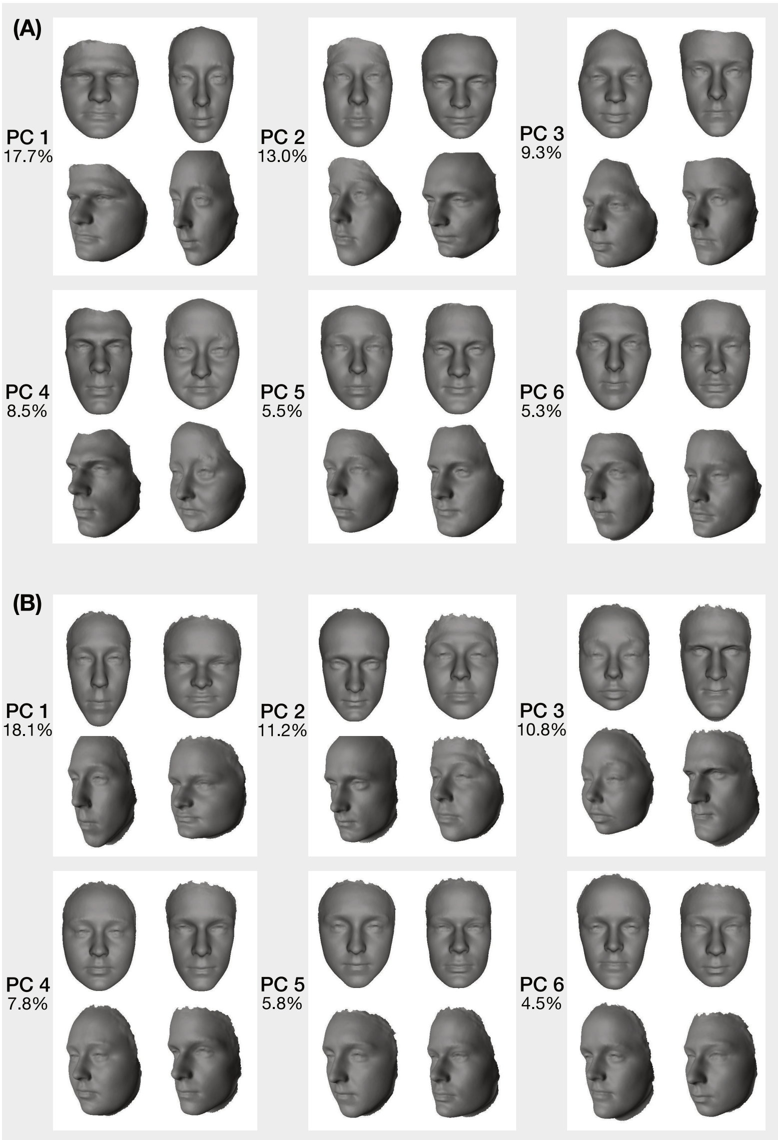


Figure C.0.1.: Principal components of shape in Face Sets 1 (A) and 2 (B), and variance explained by each principal component (PC). For each PC, shape changes associated with low (left, −5 SD) and high values (right, +5 SD) are depicted.

D. Cross-validation of face scores

In Study 1, we extended a previous measure of morphological masculinity to 3D face shape and showed that the same method can also be applied to quantify face shape correlates of height and BMI. We validated our measures perceptually by linking face-morphological scores of height and BMI to perceptual ratings of height and weight. Here, we tested whether morphological height and weight scores themselves could be cross-validated with an independent set of faces. We used Face Set 2 to construct weight and height vectors and tested whether these out-of-set vectors would produce face scores that (A) were correlated with the original, within-set derived scores, and (B) would predict perceptual ratings of height and weight.

Following the method described in Section 8.2.1.2, head models of Face Set 2 were subjected to a principal component analysis, and resulting principal component scores were used to construct height and BMI vectors accordingly (see Section 14 for a description of the prototypes used to construct the vectors). Projecting faces from Face Set 2 onto height and BMI vectors produced face-morphological height and BMI scores that were un-correlated with each other, but correlated with actual height and BMI (see Section 14.1).

In a next step, principal component scores of the 40 male and 40 female faces in the original set (Face Set 1) were predicted based on the principal component model of Face Set 2. Morphological height and weight scores were then calculated by projecting faces from Face Set 1 onto the height and BMI vectors derived from Face Set 2.

Morphological height and BMI scores as derived from Face Sets 1 (within-set) and 2 (out-of-set) were highly correlated (height: Pearson's $r(40)=.84$, $p\leq.001$); BMI: Pearson's $r(40)=.90$, $p\leq.001$). In addition, face-morphological scores derived from Face Set 2 predicted perceived height and weight (height: $\beta=.39$, $t=2.63$, $F=6.94$, $p=.012$, $R^2=.13$; weight: $\beta=.82$, $t=7.92$, $F=62.68$, $p\leq.001$, $R^2=.67$).

Scores derived from Face Set 2 also showed a trend to predict actual height ($\beta=.29$, $t=1.85$, $F=3.41$, $p=.073$, $R^2=.06$) and BMI ($\beta=.30$, $t=1.91$, $F=3.67$, $p=.063$, $R^2=.09$). Taken together, these findings suggest that the method we describe provides morphological descriptors that are stable across independent sets of faces and thus can be regarded as 'validated'.

E. The effect of stimulus dimension

Study 2 compared masculinity preferences in 2D and 3D faces. Evidence for systematic differences was found. A higher level of masculinity was consistently preferred in 3D compared to 2D faces, and some individual differences affected masculinity preferences differently in 2D and 3D faces. These preference differences might have two sources: (1) 2D and 3D stimuli are perceived differently; (2) preferences for stimulus faces that have been manipulated in their masculinity are sensitive to the specific prototypes on which masculinity transforms are based.

As stimulus faces and prototypes differed for 2D and 3D samples in Study 2, we could not directly test these possibilities. Here, we investigated whether 2D and 3D stimuli are perceived differently. We compared perceptions of masculinity and weight from 2D and 3D images of the same participants. As 3D faces might provide a more comprehensive impression of faces than 2D images do, we also tested whether accuracy in ratings might be higher for 3D compared to 2D faces by linking perceptions of weight to actual BMI. We predicted that actual BMI should be more strongly linked to perceived weight in 3D compared to 2D faces. The second alternative, i.e. that masculinity measures might be set- or prototype-sensitive, is investigated in Appendix F.

E.1. Methods and material

E.1.1. Stimulus images

For 45 men from Face Set 2, 2D images were available in addition to 3D images. To minimize potential differences between 2D and 3D stimuli, 2D images, too, were masked and rendered with an average male skin texture (see Section 8.2). Figure E.1.1 shows an example of the same man's face in 2D and 3D.

E.1.2. Face ratings

E.1.2.1. Masculinity

Sixty participants were recruited from the United States of America through Amazon MTurk and reimbursed with $2.00 each. Participants who failed to complete a brief demographic

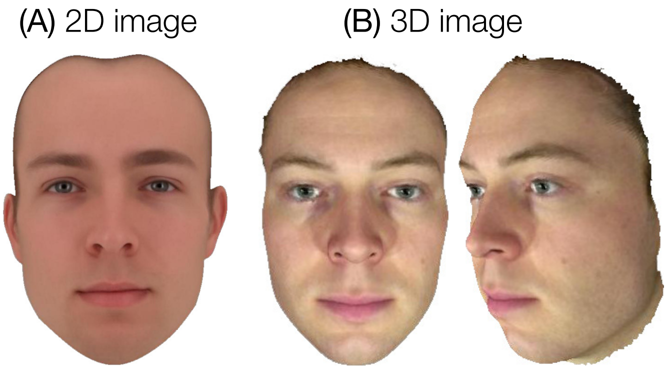


Figure E.1.1.: Example of male face from Face Set 2. (A) Masked and texture-standardised 2D image. (B) Masked and texture-standardised image captured with 3D camera.

questionnaire on their gender, age and ethnicity at the beginning of the study or who reported an ethnic background other than white were excluded. In addition, analyses were restricted to female participants[1], resulting in a final sample size of 18 women (M_{age} = 37.1±8.9 years, range 21–52 years).

Participants rated the masculinity of male 2D and 3D faces and the femininity of female 2D and 3D faces in four separate blocks (only male data is presented here). The order of blocks was pseudo-randomised, so that female faces and male faces were presented in subsequent blocks. The order of which face sex was presented first was randomised. Within each sex, the order of 2D and 3D blocks was randomised, as was the order of face stimuli within each block. Each of the four rating blocks was preceded by a movie that showed static 2D frontal image of all the faces in the respective block for one second each to provide an overview of stimulus variability. All face stimuli were presented individually against a black background and remained visible until a rating was made. Three-dimensional face stimuli were presented 'bobbing' in a sinusoidal manner from left to right and up and down. For each face, participants were asked "How masculine/feminine is this face?" Ratings were given on a slider scale beneath each image that ranged from 1-*Not masculine/feminine at all* to 100-*Very masculine/feminine* (numerical values not visible to participants).

E.1.2.2. Weight

Ratings on perceived weight were separately collected for 2D and 3D faces.[2] Participants were again recruited from the United States of America through Amazon MTurk and were

[1] Men's ratings were excluded for two reasons; first, to make findings comparable to those in Section 8; second, for men even more so than for women, perceived male masculinity might be related to perceived dominance; men's rating of male masculinity might be affected by men's self-perceived dominance (Watkins et al., 2010).
[2] Weight ratings on 2D faces were collected as part of the data collection for Study 3.4.

paid \$2.00 each. The same exclusions were applied as for masculinity ratings, except for the exclusion of male participants. Fifty participants completed the 2D rating task; after exclusions, this left 22 women and 26 men (M_{age} = 38.3±10.9 years, range 23–60). Sixty participants completed the 3D rating task; after above described exclusions, this left 21 women and 21 men (M_{age} = 40.6±11.3 years, range 23–65).

Participants rated male and female faces in two separate blocks (only male data presented here). The order of blocks and the order of face stimuli within each block was randomised. Again, face stimuli were individually presented against a black screen and remained visible until a rating was made. Three-dimensional stimuli were again presented in moving up and down and side-ways. For each face, participants were asked "How would you judge this man's/woman's weight?" Ratings were given on a slider scale beneath each image that ranged from 1-*Underweight* to 100-*Overweight* (numerical values not visible to participants).

E.2. Results

To compare masculinity perceptions in 2D and 3D faces, a linear mixed effect model with perceived masculinity as the outcome variable, stimulus dimension as the predictor and random intercepts for stimulus face nested in participant was tested. Stimulus dimension was found to have a significant effect on perceived masculinity (t=−4.76, p≤.001). Three-dimensional stimuli were rated as significantly less masculine than 2D stimuli.

To compare 2D and 3D perceptions of weight, a linear mixed effect model with perceived weight as the outcome variable, stimulus dimension as the predictor and random intercepts for stimulus face and participant was tested. Again, stimulus dimension was found to have a significant effect on perceptual ratings; 3D stimuli were rated as significantly less heavy than 2D stimuli (t=−3.93, p≤.001).

Accuracy of perceptual weight ratings from 2D and 3D faces was also tested by adding actual BMI as a predictor and allowing it to interact with stimulus dimension. The model showed that the interaction between stimulus dimension and BMI was not significant (t=−0.74, p=.459). Dropping the interaction revealed a main effect of both stimulus dimension (t=−4.18, p<.001) and BMI (t=1.74, p<.001). As can be seen from Figure E.2.1, actual BMI was a strong predictor of perceived weight. While 3D faces were rated as generally lighter than 2D faces, accuracy of 2D and 3D weight ratings did not differ.

E.3. Discussion

This follow-up study tested for differences in the perception of 2D and 3D faces using the examples of perceived masculinity and weight. We found evidence for systematic differences

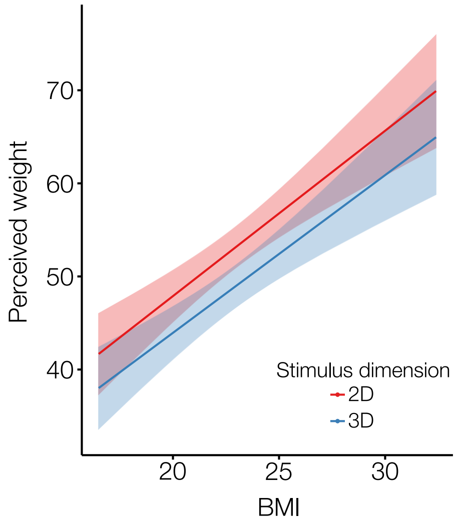


Figure E.2.1.: Accuracy of weight ratings from 2D and 3D faces did not differ. Bands represent 95% CI.

in the perception of 2D and 3D images of the same men. Men were perceived to look less masculine and less heavy when their images were presented as 3D as compared to 2D images, which might explain why Study 2 found that masculinity preferences for 3D faces were higher than those for 2D faces. Contrary to our prediction, accuracy of weight ratings was not higher for 3D compared to 2D faces. Our prediction that preferences for 3D and 2D images might differ because 3D images provide a more "accurate" representation of faces was thus not supported (or at least not so for the example of perceived weight).

F. Morphological masculinity

Study 2 found evidence for systematic differences in masculinity preferences between a set of 2D and 3D faces that had been manipulated in their facial masculinity. Appendix E tested for differences in perception due to dimensionality of stimuli. However, the two image sets that served as stimuli in Study 2 did not only differ in their dimensionality—they also differed in the prototypes that were used to manipulate facial masculinity. Here, we tested whether morphological masculinity (and thus visual manipulations of masculinity) might be to some extent set-dependent. That is, we tested whether the visual correlates of sexual dimorphism vectors used to manipulate facial images depend on the male and female prototypes they were based on.

Appendix D suggested that facial height and BMI scores are fairly stable across sets—within- and out-of-set derived scores were highly correlated, and out-of-set derived scores predicted perceptions of height and weight. Thus, we predicted that face-morphological masculinity scores derived within- and out-of-set, too, would be positively correlated and explain a similar amount of variance in perceptions of masculinity (although, as Study 1 discussed, rated masculinity is only moderately affected by sexually dimorphic shape). In addition, we tested whether within- and out-of-set face-morphological masculinity scores would show similar patterns of association with face-morphological height and BMI scores.

F.1. Methods and material

F.1.1. Perceived masculinity

The masculinity ratings described in Appendix E served as a perceptual measure of masculinity.

F.1.2. Morphological masculinity

Following the procedures outlined in Section 8.2 and Appendix D, morphological masculinity was calculated in two ways. First, by relating each face to the difference between the average male and female shape of Face Set 2 (within-set), and second, by relating the face shape of each man in Face Set 2 to the difference between the average male and female shape of Face

Set 1 (out-of-set). Section 4 provides a demographic and visual description of the average male and female faces of Face Sets 1 and 2.

F.2. Results

Morphological masculinity scores for men's faces as derived from within- and out-of-set were positively correlated ($r(50)$=.43, p=.002, see Figure F.2.1); when looking at masculinity scores for both male and female faces, this association was even stronger ($r(118)$=.77, p≤.001, see Figure F.2.2). Within-set masculinity scores correctly predicted sex in a discriminant analysis for 90.7% of faces (Wilks' λ=.321; df=1; χ^2=131.16, p<.001, comparable to findings in Section 8.2.1.2). Out-of-set masculinity scores performed slightly worse in predicting sex, but still classified 83.1% of faces correctly (Wilks' λ=.504; df=1; χ^2=79.15, p<.001).

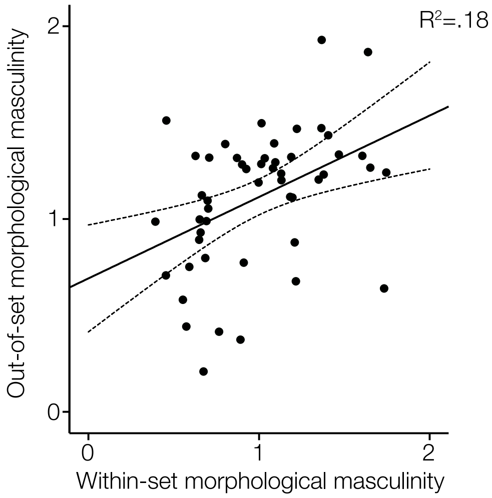


Figure F.2.1.: Within- and out-of-set derived masculinity scores were moderately correlated.

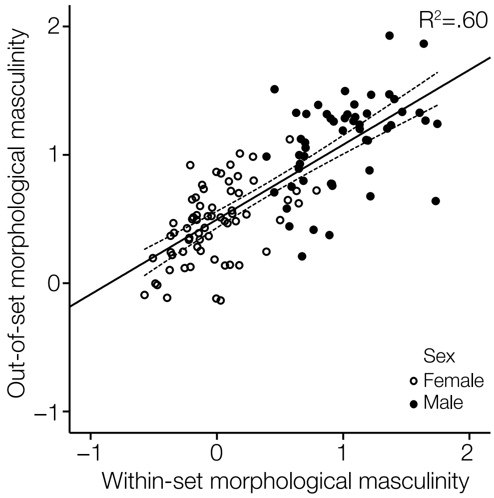


Figure F.2.2.: Across both sexes, the association of within- and out-of-set derived face-morphological masculinity scores was comparable in strength to that observed for face-morphological height and BMI scores (see Appendix D).

Despite a moderate correlation of scores for men's faces, and a strong correlation of masculinity scores when looking at both sexes, there was evidence suggesting dissimilarities between the two sets of masculinity scores. Study 1 found that morphological masculinity predicted about 11% of the variance in perceived masculinity for faces in Face Set 1. Relating within-set masculinity scores to perceived masculinity ratings of faces in Face Set 2 showed that within-set masculinity scores explained only 3% of the variance in perceived masculinity

(β=.17, F(1,48)=1.00, p=.322). In contrast, out-of-set masculinity scores explained 19% of the variance in perceived masculinity (β=.44, F(1,48)=11.47, p=.001).

To visualize similarities and differences between sexual dimorphism vectors derived from Face Sets 1 and 2, we visualized the endpoints of the masculinity range used in Study 2 based on male and female prototypes from Sets 1 and 2 (see Figure F.2.3). In addition, we compared the facial correlates of morphological masculinity scores. We created composite images of faces that scored low and high on each of the two sex dimorphism vectors by averaging faces that fell into the first and fourth quartile on each vector, and exaggerated differences between low and high masculinity composite faces by applying +200% and −200% of the difference between low and high masculinity to the average male face of Face Set 2 (see Figure F.2.4).

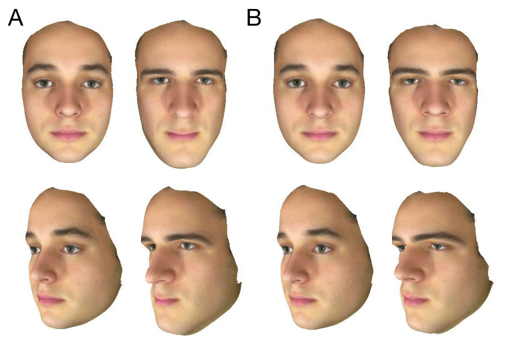


Figure F.2.3.: One of the composite faces used in Study 2 manipulated towards lower (−100%) and higher (+200%) masculinity. (A) shows masculinity manipulations based on sex prototypes from Face Set 2 (as used in Study 2), while (B) visualises masculinity manipulations based on sex prototypes from Face Set 1.

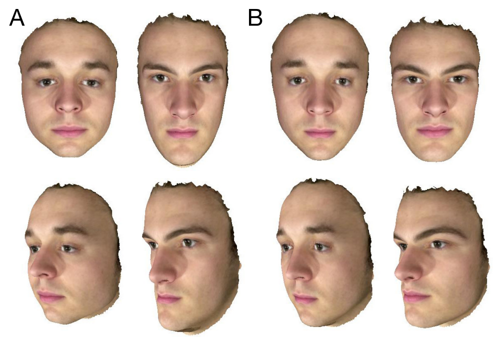


Figure F.2.4.: Shape correlates of face-morphological masculinity scores. (A) shows face shape associated with low (left) and hight (right) masculinity scores as derived from Face Set 2. (B) shows face shape associated with low (left) and high (right) masculinity scores as derived from Face Set 1.

While Figure F.2.3 shows that masculinity transformations based on Face Sets 1 and 2 show many similarities, Figure F.2.4 makes differences more salient. One of the most striking differences between masculinity based on Face Sets 1 and 2 is that morphological masculinity based on Face Set 2 appears to involve a greater change in facial elongation compared to Face Set 1. The current thesis and previous work have linked facial elongation to increased body height. Indeed, morphological masculinity scores based on Face Set 2 were found to be positively correlated with face-morphological height scores (within-set height scores: r(50)=.43, p=.002; out-of-set height scores: r(50)=.28, p=.047). Note that in contrast (and in line with findings from Study 1 of the current thesis), morphological masculinity scores derived from Face Set 1 correlated negatively with face-morphological height scores (within-set height scores: $r(50)=-.25$, $p=.080$; across-set height scores: $r(50)=-.32$, $p=.022$). Neither set of masculinity scores was significantly correlated with face-morphological BMI scores, although it should be noted that the masculinity scores derived from Set 1 showed a trend for a positive association with BMI scores ($r(50)=.23$, $p=.104$), while Set 2 scores showed a non-significant *negative* association with BMI scores ($r(50)=-.10$, p=.509).

F.3. Discussion

This follow-up study investigated whether morphological masculinity is dependent on the faces that form the female and male prototypes on which measures of masculinity (but also visual masculinity transforms) are based. Our results suggest that at least in the current sample, morphological masculinity appeared to be to some extent set-dependent, and more so than other face-morphological scores (height and BMI, see Appendix D). Morphological masculinity scores for the same faces derived from two different sets of prototypes were moderately correlated, and differed in their predictive value of perceived masculinity. Interestingly, out-of-set masculinity scores were actually a better predictor of perceived masculinity than within-set scores. The two face sets did not differ in their mean male and female height or BMI; nonetheless, our findings suggest that morphological masculinity comprised face shape correlates of sexual dimorphism in BMI and height to a different extent. It might be that while BMI did not differ between the two sets of faces, body composition, i.e. relative proportion of muscle and fat, might have differed to a different extent in Sets 1 and 2. While this finding of set-dependency might be dismissed as being related to a relatively small sample size in the current study, we note that recent studies have used similar sample sizes (e.g., Little, DeBruine & Jones, 2013; Watkins, Debruine, Feinberg & Jones, 2013, used 50 male and 50 female faces). Set-dependent masculinity is a potential methodological confound that future studies manipulating (and quantifying) facial masculinity based on linear differences between average male and female face shape should account for.

G. Ethical application forms

University of St Andrews

University Teaching and Research Ethics Committee

16 April 2012

Ethics Reference No: *Please quote this ref on all correspondence*	PS8587
Project Title:	Typically Male, Typically Female
Researcher's Name:	Iris Holzleitner
Supervisor:	Professor David Perrett

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 4th April 2012. The following documents were reviewed:

1. Ethical Application Form 13/04/2012
2. Advertisement 13/04/2012
3. Participant Information Sheet 13/04/2012
4. Consent Form 13/04/2012
5. Debriefing Form 13/04/2012
6. Questionnaire 13/04/2012

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof. D. Perrett (Supervisor) ✓
School Ethics Committee

UTREC Convenor, Mansefield, 3 St Mary's Place, St Andrews, KY16 9UY
Email: utrec@st-andrews.ac.uk Tel: 01334 462866
The University of St Andrews is a charity registered in Scotland: No SC013532

University of St Andrews

University Teaching and Research Ethics Committee

16 April 2012

Ethics Reference No: *Please quote this ref on all correspondence*	PS8588
Project Title:	What faces tell us
Researcher's Name:	Iris Holzleitner
Supervisor:	Professor David Perrett

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 4th April 2012. The following documents were reviewed:

1. Ethical Application Form 13/04/2012
2. Advertisement 13/04/2012
3. Participant Information Sheet 13/04/2012
4. Consent Form 13/04/2012
5. Debriefing Form 13/04/2012
6. Questionnaire 13/04/2012

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof. D. Perrett (Supervisor) ✓
School Ethics Committee

UTREC Convenor, Mansefield, 3 St Mary's Place, St Andrews, KY16 9UY
Email: utrec@st-andrews.ac.uk Tel: 01334 462866

University of St Andrews
from first to foremost

600 YEARS
1413 – 2013

Project Title	What Faces Tell Us
Researcher's Name	Iris Holzleitner
Supervisor	Professor David Perrett
Department/Unit	School of Psychology
Ethical Approval Code (Approval allocated to Original Application)	PS8588
Original Application Approval Date	13 April 2012
Amendment Application Approval	10 July 2012

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered at the Psychology School Ethics Committee meeting on the 10th July 2012. The following documents were reviewed:

1. Ethical Amendment Application Form 10/07/2012

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof. D. Perrett (Supervisor)
School Ethics Committee

File

University of St Andrews
from first to foremost

600 YEARS
1413 – 2013

Project Title	What Faces Tell Us
Researcher's Name	Iris Holzleitner
Supervisor	Professor David Perrett
Department/Unit	School of Psychology & Neuroscience
Ethical Approval Code (Approval allocated to Original Application)	PS8588
Original Application Approval Date	13 April 2012
Amendment Application Approval	26 September 2012

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 26th September 2012. The following documents were reviewed:

1. Ethical Amendment Application Form 26/09/2012
2. Advertisement 26/09/2012
3. Participant Information Sheet 26/09/2012
4. Questionnaire 26/09/2012

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Professor D. Perrett (Supervisor)
School Ethics Committee

University of St Andrews
from first to foremost
600 YEARS 1413 – 2013

Project Title	What Faces Tell Us
Researchers' Names	Iris Holzleitner, Maggie Webb, Aida Sincheva
Supervisor	Professor David Perrett
Department/Unit	School of Psychology & Neuroscience
Ethical Approval Code (Approval allocated to Original Application)	PS8588
Original Application Approval Date	13 April 2012
Amendment Application Approval	10 June 2013

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered by the Psychology & Neuroscience School Ethics Committee on the 10th June 2013. The following documents were reviewed:

1. Ethical Amendment Application Form 10/06/2013
2. Participant Information Sheet 10/06/2013
3. Consent Form 10/06/2013
4. Debriefing Form 10/06/2013

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission for an amendment to the original approved research proposal only. If you are unable to complete your research within the original three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant changes which occur in connection with this study, and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof. D. Perrett (Supervisor)
School Ethics Committee

University of St Andrews
from first to foremost
600 YEARS 1413 – 2013

Project Title	What Faces Tell Us
Researchers' Name	Iris Holzleitner and Shona Miller
Supervisor	Professor David Perrett
Department/Unit	School of Psychology & Neuroscience
Ethical Approval Code (Approval allocated to Original Application)	PS8588
Original Application Approval Date	13 April 2012
Amendment Application Approval	02 October 2014

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered by the Psychology & Neuroscience School Ethics Committee on the 2nd October 2014. The following documents were reviewed:

1. Ethical Amendment Application Form 02/10/2014

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Professor D. Perrett (Supervisor)

School of Psychology & Neuroscience, St Mary's Quad, South Street, St Andrews, Fife KY16 9JP
Email: psyethics@st-andrews.ac.uk Tel: 01334 462071

University of St Andrews

University Teaching and Research Ethics Committee

01 June 2012

Ethics Reference No: *Please quote this ref on all correspondence*	PS8792
Project Title:	What faces tell us – Online Study
Researcher's Name:	Iris Holzleitner
Supervisor:	Professor David Perrett

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 2nd May 2012. The following documents were reviewed:

1. Ethical Application Form 01/06/2012
2. Advertisement 01/06/2012
3. Participant Information Sheet 01/06/2012
4. Consent Form 01/06/2012
5. Debriefing Form 01/06/2012
6. Questionnaire 01/06/2012

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof. D. Perrett (Supervisor)
School Ethics Committee

UTREC Convenor, Mansefield, 3 St Mary's Place, St Andrews, KY16 9UY
Email: utrec@st-andrews.ac.uk Tel: 01334 462866

University of St Andrews
from first to foremost

600 YEARS
1413 – 2013

Project Title	What Faces Tell Us – Online Study
Researcher's Name	Iris Holzleitner
Supervisor	Professor David Perrett
Department/Unit	School of Psychology & Neuroscience
Ethical Approval Code (Approval allocated to Original Application)	PS8792
Original Application Approval Date	01 June 2012
Amendment Application Approval	19 April 2013

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered by the Psychology & Neuroscience School Ethics Committee on the 19th April 2013. The following documents were reviewed:

1. Ethical Amendment Application Form 19/04/2013

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission for an amendment to the original approved research proposal only. If you are unable to complete your research within the original three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant changes which occur in connection with this study, and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof. D. Perrett (Supervisor)
School Ethics Committee

University of St Andrews
from first to foremost

600 YEARS
1413 – 2013

Project Title	What Faces Tell Us – Online Study
Researchers' Names	Iris Holzleitner, Maggie Webb, Aida Sincheva
Supervisor	Professor David Perrett
Department/Unit	School of Psychology & Neuroscience
Ethical Approval Code (Approval allocated to Original Application)	PS8792
Original Application Approval Date	01 June 2012
Amendment Application Approval	10 June 2013

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered by the Psychology & Neuroscience School Ethics Committee on the 10th June 2013. The following documents were reviewed:

1. Ethical Amendment Application Form 10/06/2013
2. Participant Information Sheet 10/06/2013
3. Consent Form 10/06/2013
4. Debriefing Form 10/06/2013

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission for an amendment to the original approved research proposal only. If you are unable to complete your research within the original three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant changes which occur in connection with this study, and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof. D. Perrett (Supervisor)
School Ethics Committee

University of St Andrews
from first to foremost

600 YEARS
1413 – 2013

Project Title	What Faces Tell Us – Online Study
Researchers' Names	Iris Holzleitner and Kerry Harcus
Supervisor	Professor David Perrett
Department/Unit	School of Psychology & Neuroscience
Ethical Approval Code	PS8792
Original Application Approval Date	01 June 2012
Amendment Application Approval	21 February 2014

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered by the Psychology & Neuroscience School Ethics Committee on the 21st February 2014. The following documents were reviewed:

1. Ethical Amendment Application Form 21/02/2014
2. Participant Information Sheet 21/02/2014
3. Consent Form 21/02/2014
4. Debriefing Form 21/02/2014
5.

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof D. Perrett (Supervisor)
School Ethics Committee

School of Psychology & Neuroscience, St Mary's Quad, South Street, St Andrews, Fife KY16 9JP
Email: psyethics@st-andrews.ac.uk Tel: 01334 462071

University of St Andrews
from first to foremost

600 YEARS
1413 – 2013

Project Title	What Faces Tell Us – Online Study
Researcher's Name	Iris Holzleitner
Supervisor	Professor David Perrett
Department/Unit	School of Psychology & Neuroscience
Ethical Approval Code (Approval allocated to Original Application)	PS8792
Original Application Approval Date	01 June 2012
Amendment Application Approval	10 October 2014

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 7th October 2014. The following documents were reviewed:

1. Ethical Amendment Application Form 10/10/2014
2. Participant Information Sheet 10/10/2014
3. Consent Form 10/10/2014
4. Debriefing Form 10/10/2014

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Professor David Perrett (Supervisor)
School Ethics Committee

University of St Andrews

University Teaching and Research Ethics Committee

26 September 2012

Ethics Reference No: *Please quote this ref on all correspondence*	PS9157
Project Title:	Does Size Matter?
Researcher's Name:	Iris Holzleitner
Supervisor:	Professor David Perrett

Thank you for submitting your application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 19th September 2012. The following documents were reviewed:

1. Ethical Application Form 25/09/2012
2. Advertisement 25/09/2012
3. Participant Information Sheet 25/09/2012
4. Consent Form 25/09/2012
5. Debriefing Form 25/09/2012
6. Questionnaire 25/09/2012

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof. D. Perrett (Supervisor)
School Ethics Committee

UTREC Convenor, Mansefield, 3 St Mary's Place, St Andrews, KY16 9UY
Email: utrec@st-andrews.ac.uk Tel: 01334 462866
The University of St Andrews is a charity registered in Scotland: No SC013532

University of St Andrews
from first to foremost

600 YEARS
1413 – 2013

Project Title	Does Size Matter?
Researcher's Name	Iris Holzleitner
Supervisor	Professor David Perrett
Department/Unit	School of Psychology & Neuroscience
Ethical Approval Code (Approval allocated to Original Application)	PS9157
Original Application Approval Date	25 September 2012
Amendment Application Approval	7 November 2012

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered by the Psychology & Neuroscience School Ethics Committee on the 7th November 2012. The following documents were reviewed:

1. Ethical Amendment Application Form 07/11/2012
2. Advertisement 07/11/2012
3. Participant Information Sheet 07/11/2012
4. Consent Form 07/11/2012
5. Questionnaire 07/11/2012

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof D, Perrett (Supervisor) ✓
School Ethics Committee

Note:
Revised Approval Code (hence new copy of original approval)

University of St Andrews
from first to foremost

600 YEARS
1413 – 2013

Project Title	Does Size Matter?
Researchers Name	Iris Holzleitner
Supervisor	Professor David Perrett
Department/Unit	School of Psychology & Neuroscience
Ethical Approval Code (Approval allocated to Original Application)	PS9157
Original Application Approval Date	25 September 2012
Amendment Application Approval	19 April 2013

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered by the Psychology & Neuroscience School Ethics Committee on the 19th April 2013. The following documents were reviewed:

1. Ethical Amendment Application Form 19/04/2013
2. Advertisement 19/04/2013
3. Participant Information Sheet 19/04/2013
4. Consent Form 19/04/2013
5. Questionnaire 19/04/2013

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission for an amendment to the original approved research proposal only. If you are unable to complete your research within the original three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant changes which occur in connection with this study, and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof. D. Perrett (Supervisor)
School Ethics Committee

University of St Andrews

University Teaching and Research Ethics Committee
Sub-committee

27th March 2014

Ethics Reference No: *Please quote this ref on all correspondence*	PS10919
Project Title:	In good shape and colour
Researchers' Names:	Prof. David Perrett, Iris Holzleitner, Audrey Henderson, Carlota Batres
Supervisor:	Prof. David Perrett

Thank you for submitting your application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 26th of March 2014. The following documents were reviewed:

1. Ethical Application Form 26/03/14
2. Participant Information Sheet 26/03/14
3. Consent Form 26/03/14
4. Debriefing Form 26/03/14
5. Questionnaires 26/03/14

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' https://www.st-andrews.ac.uk/utrec/guidelines/ are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof. David Perrett (Supervisor)
School Ethics Committee

School of Psychology & Neuroscience, St Mary's Quad, South Street, St Andrews, Fife KY16 9JP
Email: psyethics@st-andrews.ac.uk Tel: 01334 462071

University of St Andrews
from first to foremost

600 YEARS
1413 – 2013

Project Title	In good shape and colour
Researchers Names'	Iris Holzleitner, Audrey Henderson and Carlota Batres
Supervisor	Professor David Perrett
Department/Unit	School of Psychology & Neuroscience
Ethical Approval Code (Approval allocated to Original Application)	PS10919
Original Application Approval Date	26 March 2014
Amendment Application Approval	26 May 2015

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered by the Psychology & Neuroscience School Ethics Committee on the 26th May 2015. The following documents were reviewed:

1. Ethical Amendment Application Form 26/05/2015
2. Participant Information Sheet 26/05/2015
3. Consent Form 26/05/2015
4. Debriefing Form 26/05/2015

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof D Perrett (Supervisor)
School Ethics Committee

School of Psychology & Neuroscience, St Mary's Quad, South Street, St Andrews, Fife KY16 9JP
Email: psyethics@st-andrews.ac.uk Tel: 01334 462071