Bachelor Degree Project

UNIVERSITY OF SKÖVDE

THREATENING MEASURES, AT FACE VALUE:
Electrophysiology Indicating Confounds of the Facial Width-to-Height Ratio

Bachelor Degree Project in Cognitive Neuroscience
Basic-level, 22.5 ECTS
Spring term 2019

Carl Lindersson

Supervisor: Oskar MacGregor
Examiner: Antti Revonsuo

Abstract

Previous studies support that the relative width of the upper face (facial width-to-height ratio; fWHR) has evolved to signal threat, but these studies rely greatly on subjective facial ratings and measurements prone to confounds. The present study objectively quantifies threat perception to the magnitude of the observers' electrophysiological reaction, specifically the event-related potential (ERP) called the late positive potential (LPP), and investigate if brow height and jaw width could have confounded previous fWHR studies. Swedish and international students (N = 30, females = 11, M_{age} = 24 years, SD_{age} = 2.9) were shown computer-generated neutral faces created with the underlying skeletal morphology varying in brow ridge height, cheekbone width and jaw width. Participants first rated how threatening each face was and then viewed 12 blocks of 64 faces while their electroencephalography (EEG) was recorded. The results supported that the LPP could be used to index threat perception and showed that only brow height significantly affected both facial ratings ($p < .001$, $\eta_p^2 = .698$) and magnitude of the LPP within the 400 to 800 ms latency ($p = .02$, $d = .542$). Hence, brow height, not facial width, could explain previous findings. The results contradict the hypothesis that fWHR is an evolved cue of threat and instead support the overgeneralisation hypothesis in that faces with similar features to anger will be perceived as more threatening.

Keywords: threat perception, evolutionary cues, fWHR, LPP, overgeneralisation hypothesis

Table of Content

1. Introduction

You find yourself walking along a dirt road, soaking in the view of the fields spanning either side. Suddenly, in your periphery, you spot something looking like a snake on the side of the road. In a split second, you halt to a stop and focus your gaze. On closer inspection, you see that it was only a twig laying on the road. Your heart starts beating again and you move on.

Mistaking twigs for snakes is probably common practice for all humans. But why? One proposition is that snakes have been a persistent threat to humans throughout evolution and this has put selective pressure on the ability to detect snakes (Isbell, 2006; Öhman, Flykt, & Esteves, 2001; Soares, Lindström, Esteves, & Öhman, 2014). Hence, humans would have evolved to be sensitive to cues related to snakes so that we could quickly spot one and react appropriately. That is, in absolute panic.

But not only snakes have been constant threats to humans throughout our evolutionary history. Other humans with bad intentions most likely were a greater threat than snakes (Manson & Wrangham, 1991). Hence, the same line of reasoning applied to snakes has been made for how we perceive and react to faces of others (Sell et al., 2009; Öhman, Lundqvist, & Esteves, 2001). That is, humans could have evolved a perceptual sensitivity to cues related to threatening intentions or behavioural tendencies of others. Interestingly, it has been repeatedly shown that humans not only quickly infer personality traits from expressionless faces (e.g., Geniole, Denson, Dixson, Carré, & McCormick, 2015; Haselhuhn, Ormiston, & Wong, 2015; Oosterhof & Todorov, 2008; Ormiston, Wong, & Haselhuhn, 2017) but recent studies also support that some of these inferences could be accurate (e.g.,Geniole et al., 2015; Haselhuhn et al., 2015; Sell et al., 2009). It becomes more complicated with these inferences also predicting political elections (Todorov, Mandisodza, Goren, & Hall, 2005) and death penalty sentences (Wilson & Rule, 2015). Therefore, understanding the perceptual mechanisms underlying these inferences could help identify biases in our justice systems and also have global political implications.

The facial width-to-height ratio (fWHR), a facial measure calculated as the width of the cheekbones divided by the height of the upper face, has consistently predicted both observers' perceptions of others and the behaviour of the ones observed. Specifically, men with high compared to low fWHR have been found to both be perceived as, and behave, more threatening, aggressive, dominant and untrustworthy (for meta-analyses see Geniole et al., 2015; Haselhuhn et al., 2015) Furthermore, there is support for this relationship to be sexually dimorphic with men on average having higher fWHR than women (Geniole et al., 2015; Weston, Friday, & Liò, 2007). As this sexual dimorphism could indicate that sexual selection pressures have driven the development of the fWHR, e.g., like male deer developing antlers to fight other male deer for access to females, it has been proposed that male-male competition drove the development of fWHR in men as a signal of aggression and dominance. That is, humans might have evolved perceptual mechanisms for detecting fWHR to quickly assess how aggressive and consequently how much of a threat a male might be (Carré & Mccormick, 2008; Carré, Mccormick, & Mondloch, 2009; Geniole et al., 2015; Haselhuhn et al., 2015).

However, whether fWHR truly is sexually dimorphic has been questioned due to contradictory findings of fWHR not being sexually dimorphic (Gómez-Valdés et al., 2013; Kramer, Jones, & Ward, 2012; Özener, 2012), not related to aggression (Gómez-Valdés et al., 2013; Özener, 2012) or explained by ethnic and geographical differences in skulls and faces (Kramer, 2017). There is also good reason to believe that the measurements used in most previous studies on fWHR are prone to be confounded by effects of brow height as most studies on fWHR incorporate brow height in their facial measures (e.g., Carré & Mccormick, 2008; Carré et al., 2009; Geniole et al., 2015; Ormiston et al., 2017; but see Haselhuhn et al., 2015). Additionally, other sexually dimorphic features could confound the fWHR and as the protrusion of the brow ridges, cheekbone width and jaw width are sexually dimorphic (Enlow & Hans, 1996; Fink et al., 2005; Verdonck, Gaethofs, Carels, & DeZegher, 1999), these features could also have affected previous studies on fWHR.

Furthermore, in accordance with the overgeneralisation hypothesis, faces with high fWHR might simply share similar features with angry faces, e.g., low brows, and therefore be mistaken for angry, hence threatening, faces (Carré & Mccormick, 2008; Knutson, 1996; Said, Sebe, & Todorov, 2009).

Another problem with previous studies on fWHR is that threat perception is measured through subjective ratings which opens up for potential bias, as the raters could modulate their answers if they find their initial response not socially acceptable (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). However, there might be a way to objectively quantify the observers' threat perception through measuring their electroencephalography (EEG).

By using EEG, Schupp, Öhman, et al. (2004) found that the amplitude of the event-related potential (ERP) called the late positive potential (LPP) correlated with viewing threatening faces expressing anger. Erotica and threat-related content like angry faces and mutilated limbs also elicit higher LPP magnitude linking it to intrinsically motivating content of evolutionary significance (e.g., Schupp et al., 2000; Schupp, Öhman, et al., 2004; Weinberg & Hajcak, 2010). Interestingly, Eldblom (2018) found that subjects expressed greater LPP while viewing pictures of males with higher fWHR compared to when viewing pictures of males having low fWHR. The subjects also scored the males with high fWHR as looking more aggressive. Hence, LPP seems to be able to accurately indicate threat perception related to fWHR. However, some inconsistencies in the result weaken these findings (Eldblom, 2018).

More research is needed to establish whether LPP can be used to index threat perception. Another problem is that it is currently impossible to say what makes a high fWHR be perceived as threatening. Is it due to a high fWHR being an honest cue of aggressive and dominant tendencies? In that case, are other masculine features such as protruding brow ridges and wide jaw affecting threat perception? Alternatively, are faces with high fWHR confounded by low brow height, making the faces similar to the facial expression of anger and thereby making them look threatening? The present study will address these questions by recording participants EEG while they observe neutral male faces with skeletal morphology varying brow ridge height (high vs. low), cheekbone width (slim vs. wide) and jaw width (slim vs. wide). This allows for comparisons between participants' electrophysiological responses and their subjective ratings, providing good grounds to evaluate the influence of brow height, cheekbone width and jaw width, by themselves and in combination.

Due to these features' connection to perception of threat, masculinity and dominance (Geniole et al., 2015; Said et al., 2009; Windhager, Schaefer, & Fink, 2011), the primary hypothesis is that faces with either low brows, and/or wide cheekbones and/or wide jaw will be rated as more threatening and elicit greater LPP magnitudes. The secondary hypothesis is that low brows will generate the greatest effects due to their similarities with angry faces.

To give more depth to the experiment, a broader background will explore related research and theories in more detail. Due to that the history of inferring behavioural traits from physical appearance has been suggested to shape research related to it (Valla, Ceci, & Williams, 2011), a short introduction to related historical influences will be given as a starting point. Secondly, evolutionary theories and their relation to fWHR will be considered. Thirdly, support for fWHR being an evolved cue of threat will be described. Forth, the critique of the fWHR as a valid measurement will be explored together with other explanations for its links to threat perception. Lastly, the suitability of the LPP as an objective measure of threat perception will be investigated, followed by the rationale for the study.

2. Background

2.1 Short Historical Introduction

The interest in inferring behavioural traits from appearance goes far back in history. Aristotle promoted the idea that accurate inferences could be made about a person's character, solely from observing their physical appearance (Aristotle, trans. 1984). From the ideas by the Greek philosopher, the pseudoscience called physiognomy developed with the goal of inferring behavioural dispositions from facial features. Physiognomy gained great popularity through the works of Johann Caspar Lavater during the late 18th century. However, after the unsound scientific vigour of physiognomy was illuminated during the 19th century, partly due to Cesare Lombroso's so-called criminal anthropology, it fell out of favour (for historical analysis see Twine, 2002; Zebrowitz, 1997). In 1872, Darwin published his book *The Expression of the Emotions in Man and Animals* suggesting that quick and accurate inferences of characteristics from physical signs would be of evolutionary relevance. This later sparked new life into the idea of accurate appearance-based inferences (Valla et al., 2011).

Sadly, Darwin's theories were misinterpreted and used to promote unethical policies by the Social Darwinism and eugenics movements (Valla et al., 2011). After the fall of these movements, the idea that physical appearance could accurately reveal character traits was still associated with the unethical conducts by Social Darwinism, eugenics and not to mention the Holocaust. This stigma affected the scientific community, making researchers deliberately avoid questions regarding any potential accuracy of appearance-based inferences, and instead focus on the negative consequences of stereotypes. To not judge the book by its cover became the preferred way of thinking (Valla et al., 2011).

Nevertheless, there were those who questioned whether appearance truly did not say anything about character. Berscheid (1981) appealed to the naturalistic fallacy, arguing that just because society thinks that physical appearance should not matter, does not mean that it does not matter. During the mid-1980s and early 1990s, more and more researchers began investigating Darwin's idea of adaptive accurate appearance-based inferences (Valla et al., 2011). Even though Darwin's ideas today have become separated from their prior misuse, Valla et al. (2011) suggest that an echo of the related stigma is seen in that some researchers still overlook the possibility of that physical cues could accurately infer behaviour to some degree. However, it should be mentioned that during the 21st Century more research has considered accuracy of inference-based judgements (e.g., Geniole et al., 2015; Třebický, Havlíček, Roberts, Little, & Kleisner, 2013; Valla et al., 2011), although, it has not been without reasonable scepticism (for review see Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). Regardless, historical forces might still affect reasoning within this field to some degree. Acknowledging such forces could perhaps minimise any bias that comes along with it.

2.2 Evolutionary Forces Explaining Facial Features

Darwin's (1872) idea of adaptive, quick and accurate appearance-based inference have influenced many theories within evolutionary psychology related to facial features such as fWHR. Being able to quickly infer the behavioural disposition of another person could be adaptive in several ways. Sexual selection, the evolutionary mechanism favouring traits advantageous for accessing mates, could have driven the development of specific features, e.g., facial features (Darwin, 1871). This could lead to individuals developing features attractive to the opposite sex (intersexual selection), as well as features giving the individual an advantage against their same-sex peers in the competition for access to the opposite sex (intrasexual selection) (Darwin, 1871). In most species, including humans, females invest more time and resources into their offspring than their male partner, Trivers (1985) suggested this would explain why females are pickier when choosing mates. This was in line with suggestions by Darwin (1871), that intersexual selection probably was mostly shaped by females choice of partners while, consequently, intrasexual selection was mostly shaped by male-male competition for females. After Darwin (1871) and Trivers (1985) suggestions, there has been substantial empirical support for female choice (e.g., Andersson, 1982; Haines & Gould, 1994) and male-male competition (e.g., Clutton-Brock, Guinness, & Albon, 1982; Clutton-Brock & Harvey, 1977).

As sexual selection acts differently on males and females it would explain why males and females within the same species can differ physically and behaviourally, a phenomenon called sexual dimorphism (Darwin, 1871). For example, males' typically greater muscle mass and higher tendency for aggression compared to females is often taken to be an evolutionary consequence of male-male competition (Clutton-Brock et al., 1982). Interestingly, research has found fWHR to differ between males and females, but most importantly, it was found that differences in fWHR were not a consequence of the sexually dimorphic differences between males' and females' body size (Weston et al., 2007). Specifically, it was found that the length of the males' upper faces was un-proportionately short in relation to its width. Hence, it was proposed that the difference in fWHR observed between the sexes might be a consequence of sexual selection unrelated to body size (Weston et al., 2007). Could females maybe prefer males with a relatively shorter upper face, and consequently a higher fWHR?

2.2.1 Female choice. Given the apparent evolutionary influence of female choice, one could wonder why females find specific features in men attractive. Several researchers have attempted to answer this question. Fisher (1930) proposed what he called the runaway hypothesis which argued that females prefer males with features advantageous for survival, but as a consequence, these features could over several generations evolve beyond their functional use and thereby instead become disadvantageous for survival. Hence, according to Fisher (1930), females' preference is originally tied to genetic advantage but females will prefer "good-looking" features over advantageous features. Another idea, the handicap hypothesis, was presented by Zahavi (1975) and argued that females find males with pronounced sexually dimorphic ornaments (attractive, but not necessarily functional features) attractive because these ornaments are costly to produce, maintain, and they decrease chances for survival. Hence, males surviving despite pronounced ornaments as handicaps would show they possessed attractive qualities superior to survival, indirectly making the ornaments signals of quality (Folstad & Karter, 1992; Zahavi, 1975). Furthermore, Hamilton and Zuk (1982) proposed the parasite hypothesis arguing that females prefer males with features displaying that they are free from parasites and hence have high parasite resistance. Such features would act as honest signals for genetic quality, meaning that attractive features need to correlate with actual genetic quality, and not merely be good-looking features coming at a cost of being disadvantageous.

Regarding female choice and facial features, there is support for that males' sexually dimorphic facial features act as both honest signals for genetic quality (Rhodes, 2006; Rhodes, Chan, Zebrowitz, & Simmons, 2003; Thornhill & Gangestad, 1993, 1999) and as costly ornaments displaying survival qualities (Folstad & Karter, 1992; Thornhill & Gangestad, 1993, 1999). However, females do not always rate masculine male faces as attractive (Rhodes, 2006). Although, the differences between studies have been suggested to be a result of measurement issues, therefore, there might still be a link between facial masculinity and female choice (Rhodes, 2006).

Sexually dimorphic facial features largely develop during puberty and might partly act as honest signals for sexual maturity and reproductive potential (Enlow & Hans, 1996; Fink et al., 2005; Verdonck et al., 1999). The growth of sexually dimorphic facial features during puberty in males are driven by testosterone, resulting in growth of the jaw, cheekbones, brow ridges, the length from the brows to the bottom of the nose, and facial hair (Enlow & Hans, 1996; Fink et al., 2005; Verdonck et al., 1999). Interestingly, as testosterone has been found to reduce the function of the immune system, exhibiting large masculine facial features could act as an honest signal of genetic quality, due to that only males that could handle such stress to the immune system could develop such masculine features (Folstad & Karter, 1992; Thornhill & Gangestad, 1993, 1999). This is called the immunocompetence-handicap hypothesis (Folstad & Karter, 1992). It is in line with both the handicap hypothesis and the parasite hypothesis, as developing pronounced testosterone-driven features is both a handicap and shows superior resistance to parasites even under stress to the immune system. Hence, testosterone-driven features would be honest signals of parasite burden and resistance (Folstad & Karter, 1992).

Giving further support to sexually dimorphic male features acting as honest signals of quality, Rhodes et al. (2003) found ratings of facial masculinity in adolescent males correlated with these men's health. Moreover, female perception of masculine facial features in males have also been linked to dominance (Fink, Neave, & Seydel, 2007; Mazur & Mueller, 1996) and strength

(Fink et al., 2007) which both are traits found attractive by females (Buss, 1989; Cunningham, Barbee, & Pike, 1990; Fink et al., 2007). Furthermore, males with higher sexually dimorphic faces were found to have greater reproductive success in short-term relationships (Rhodes, Simmons, & Peters, 2005). With relevance for the current study, high fWHR is a male sexually dimorphic feature and can, therefore, be regarded as linked with the above-stated findings regarding facial masculinity.

2.2.2 Male-male competition. Males with more masculine faces, and consequently relatively high fWHR, might have had an advantage in male-male competition for females. Males with more pronounced masculine facial features have been found both be perceived as, and to be, more dominant (e.g., Mazur & Mueller, 1996; Watkins, Jones, & DeBruine, 2010), physically stronger (Fink et al., 2007) and better fighters (Little, Třebický, Havlíček, Roberts, & Kleisner, 2015; Zilioli et al., 2015) compared to males with less pronounced masculine facial features. These findings suggest that high facial masculinity could signal physical formidability, that is, that the individual would be a tough opponent in a physical fight. Hence, masculine facial features would be advantageous in male-male competition as other males would be less prone to engage in physical fights due to a higher risk of injury (Sell et al., 2009; Parker, 1974).

Moreover, there are reasons to believe masculine facial features not only act as a signal of formidability but could also have direct practical functions. Given that male-male fighting has been substantial throughout evolutionary history and more usual than female fighting (Archer, 2004, 2009; Manson & Wrangham, 1991), selection pressures likely favoured males with features that increased fighting ability (Sell, Hone, & Pound, 2012). Consequently, such features would increase a male's absolute probability of winning a physical fight over a resource, that is, the male's resource-holding potential, which would affect access to females (Sell et al., 2009; Parker, 1974). Furthermore, as fights often involve attacks to the face (Shepherd, Gayford, Leslie, & Scully, 1988), there would be a greater need for males to protect their faces. Such a need may have driven the development of the more robust facial features in males (Puts, Jones, & DeBruine, 2012). In support of this suggestion, males seem to have fewer facial fractures than women (Shepherd et al., 1988).

Given masculine facial features' seemingly protective properties as well as these features' link to dominance and strength, one would expect facial masculinity to also indicate fighting ability. This has indeed been found to be the case (Little et al., 2015; Zilioli et al., 2015). By comparing the faces of fighters in The Ultimate Fighting Championship (UFC), the world largest mixed-martial art organisation, Zilioli et al. (2015) found that fighters with a relatively high fWHR won more fights than males with lower fWHR. Another study, perhaps more related to conflict outcome than fighting ability per se, investigated fWHR of skulls in relation to the cause of death in male homicide cases (Stirrat, Stulp, & Pollet, 2012). It was found that males with slimmer faces, opposite to males with wider faces, more often were killed by direct physical contact, measured as being beaten, stabbed or strangled to death. Stirrat et al. (2012) suggested that their results support that males with a higher fWHR could have had an evolutionary advantage in male-male competition. Furthermore, Stirrat et al., (2012) propose that males with high fWHR might exhibit their documented high aggressive tendencies (Carré, & Mccormick, 2008) due to their lower risk of death from direct violence. Might these male-male competition advantages also mean that males with high fWHR are perceived as more threatening than males with lower fWHR, and do men with relatively high fWHR truly exhibit more threatening behaviour?

2.3 Is fWHR an Honest Signal for Threat?

In 2015, as a growing amount of research connected high fWHR in men with aggressive behaviour and being perceived as more threatening, it was an opportune time to conduct meta-analyses investigating these relationships. This opportunity was seised by both Haselhuhn et al. (2015) and Geniole et al. (2015).

2.3.1 An honest signal for aggression? Haselhuhn et al. (2015) set out to answer whether a high fWHR in men predicts aggressive behaviour. In earlier studies, connections have been found between men's fWHR and direct aggressive behaviour (e.g., Carré, & Mccormick, 2008) as well as indirect indications of aggression, e.g., proving untrustworthy in an economic game (Stirrat & Perrett, 2010). However, as some studies did not find a significant relationship between men's fWHR and aggression, the possibility of the positive findings being due to type-1 error needed to be investigated through a meta-analysis. Haselhuhn et al. (2015) performed such a meta-analysis on 19 studies measuring both direct and indirect aggression in relation to fWHR in men and found a significant, but weak, positive correlation between fWHR and aggression ($r = .11$, $p < 001$). The correlation persisted even when only studies on direct aggression ($r = .11$, $p < 001$) or indirect aggression ($r = .11$, $p < 001$) were considered by themselves. These findings support that fWHR could predict aggressive behaviour and hence act as an honest signal for aggression. However, as the effect size was weak, the findings also showed that factors possibly unrelated to fWHR greatly moderate men's aggressive tendencies (Haselhuhn et al., 2015).

2.3.2 An honest signal for threat and dominance? While Haselhuhn et al. (2015) focused on the correlation between men's fWHR and aggression, Geniole et al. (2015) had a broader scope in their meta-analysis including both behaviour and perception related to threat, masculinity and dominance, in both men and women. Specifically, Geniole et al. (2015) aimed to answer three questions: (1) whether fWHR is positively correlated with threatening and dominant behaviour, (2) whether men's and women's fWHR affect how threatening and dominant they are perceived by others, and (3) whether fWHR differed significantly between the sexes,. To accomplish these goals Geniole et al. (2015) considered all peer-reviewed articles on fWHR found by searching the term "facial Width-Height-Ratio" in Google Scholar. Additionally, all articles citing Weston et al. (2007), the first published article to use fWHR as a measure, was also considered to be included in the analysis. This resulted in 56 articles being used in the analysis as they met the inclusion criteria of involving fWHR measures on human faces with neutral facial expressions.

2.3.2.1 *Behavioural correlations.* The results of the analysis by Geniole et al. (2015) confirmed that fWHR predicted threatening behaviour ($\bar{r} = .16, p < .001$) and that this correlation was only found in men. Threatening behaviour was measured as either behavioural or self-reported aggressive, pejorative or selfish behaviour. There were no moderating effects found by any specific form of threatening behaviour or whether the type of measure was behavioural or self-reported. Age did not modulate the correlation between fWHR and behaviour ($p = .42$). Furthermore, dominant behaviour was also found correlating with fWHR in both men and women ($\bar{r} = .12, p < .001$). Additionally, athletic sports performance was included as another indicator of dominance and found to correlate with fWHR in men ($\bar{r} = .15, p < .001$).

2.3.2.2 *Perceptual correlations.* Moreover, Geniole et al. (2015) found that higher fWHR correlated with being perceived as more threatening ($\bar{r} = .46, p < .001$) and that age modulated the effect with young faces (age < 25) being perceived as more threatening than older faces (age > 25). Perception of dominance also correlated with fWHR ($\bar{r} = .20, p = .007$) with stronger correlation when only male faces were used ($\bar{r} = .30, p < .001$). Hence, a high fWHR not only seemed to correlate with behavioural tendencies but also act as a perceptual cue of threat and dominance (Geniole et al., 2015).

2.3.3 Support for fWHR being sexually dimorphic. Whether high fWHR in men, seemingly acting as a cue of threat, was a product of male-male competition or female choice was considered by Geniole et al. (2015) in their discussion. In their study (Geniole et al., 2015), fWHR was indeed found to be sexually dimorphic with men having a higher fWHR than women ($\bar{d} = .11, p = .009$). Furthermore, the result showed that high fWHR correlated with the face being perceived as more masculine and that the effect was only significant when the stimuli included male faces ($\bar{r} = .30, p < .001$) and stronger when including only male faces ($\bar{r} = .35, p < .001$). However, Geniole et al. (2015) point out that studies including faces of women as stimuli were few which limit how well these results account for perception of masculinity in women's faces. Nonetheless, the evidence point towards fWHR being sexually dimorphic and predicting perception of masculinity in men only. Interestingly, Geniole et al. (2015) found that when observers consisted mostly out of women, men with high fWHR was rated as less attractive than men with lower fWHR. Hence it was suggested that fWHR probably mainly play a part in male-male competition and not female choice (Geniole et al., 2015).

More support for that male-male competition drove fWHR to be sexually dimorphic was found in additional analyses by Geniole et al. (2015). One relevant finding was that the correlation between fWHR and threat perception of young (age < 25) faces was particularly strong for perception of aggression in relation to the other types of threatening characteristics, e.g., untrustworthiness. Furthermore, young faces with high fWHR also elicited particularly strong perceptions of threat. Discussing these above-mentioned results, Geniole et al. (2015) propose that the results might be due to that young males are the most prone to commit violent acts, as described by Wilson and Daly (1985). Hence, fWHR might foremost act as a cue of aggression and the age of the observed face might further influence the strength of the cue, confounding the sole effect by fWHR. Furthermore, in their analysis, Geniole et al. (2015) also found fWHR to be significantly positively correlated with athletic performance ($\bar{r} = .10$, $p = .04$) and business outcome, e.g., negotiation ability ($\bar{r} = .32$, $p = .002$). Hence, males with a higher fWHR seemed to have advantages in competition with males with lower fWHR. Altogether, these findings were taken to suggest that the sexual dimorphism of the fWHR has been formed by male-male competition to act as a cue of threat and dominance (Geniole et al., 2015).

2.4 Findings Against the fWHR Being a Cue of Threat.

As noted in Geniole et al. (2015), fWHR was not the only factor to affect how threatening observers found male faces, e.g., the age of the faces also seemed to influence the effect size. Apart from age, might there also be other factors, not directly accounted for by the fWHR measure, that could influence how threatening a face is perceived? Such potential facial features would, if not controlled for, confound the effect of the fWHR has on observers threat perceptions.

2.4.1 Features related to body size and weight. As mentioned earlier, the fWHR was first mentioned in Weston et al. (2007) as a sexually dimorphic facial feature that was greater in males than females and independent of changes in body size. This was then taken to indicate that fWHR, in contrast to other facial features, could have evolved due to sexual selection pressures on specific facial features, unrelated to selective pressures of body size (Weston et al., 2007). Carré & Mccormick (2008) then found that fWHR correlated with aggressive behaviour in men and argued it suggested that fWHR could be an honest cue for aggression. From there on, research connecting fWHR to threat perception started to be published (see Geniole et al., 2015). Interestingly, Geniole et al. (2015) found that fWHR predicted body size ($\overline{r}$ = .31, p < .001) which suggests that the behavioural tendencies and perceptual judgements related to fWHR could be related to body size after all. Worth mentioning is that there are studies that have controlled for body mass index (BMI), as Zilioli et al. (2015) who found a positive correlation between fWHR and perception of MMA fighters' formidability (r = .460, p = .001, n = 32) and that this correlation persisted after controlling for BMI (r = .338, p = .02, n = 32). However, as the strength of the correlation dropped when controlling for BMI, its seems like BMI might have influenced perceptions of formidability to some degree after all.

Another study that nicely complements the findings by Zilioli et al. (2015) is Třebický et al. (2013) who also investigates the link between facial features and perception of MMA fighters' fighting ability while controlling for the fighters' weight. Interestingly, Třebický et al. (2013) found that weight predicted both perceived aggressiveness (R^2 = .089, p < .001) and perceived fighting ability (R^2 = .085, p < .001). However, using a geometric morphometric technique, Třebický et al. (2013) identified a set of facial features that predicted perceived aggressiveness and interacted with weight (see figure 1). This showed that the facial features that mediated perception of aggressiveness were associated with the weight of the fighters (Třebický et al., 2013). Given that body weight has been linked to both aggressiveness (Deaner, Goetz, Shattuck, & Schnotala, 2012) and strength (Sell et al., 2009) it is reasonable to believe facial features cueing weight would be of importance in male-male competition, at least in athletic individuals as MMA fighters.

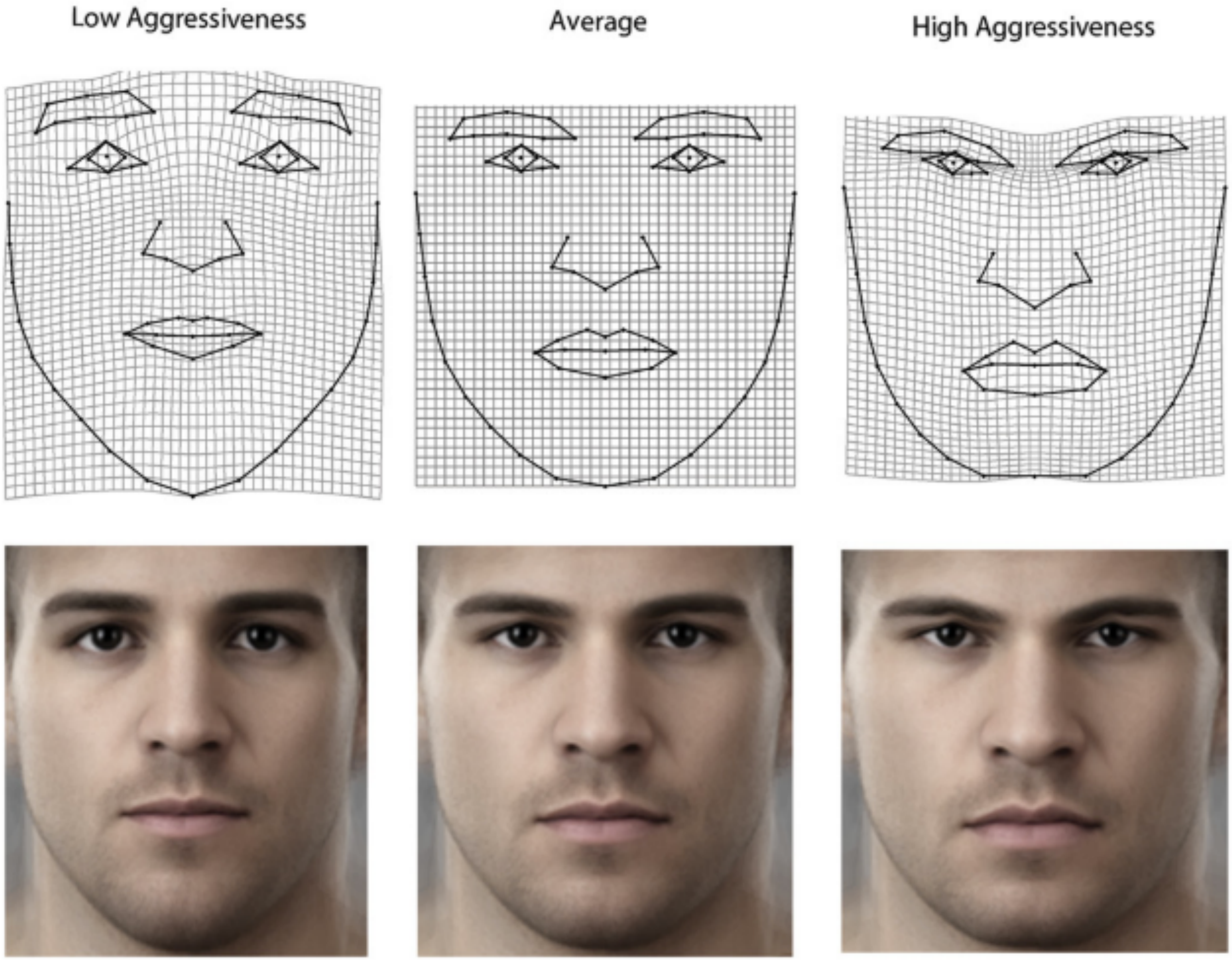

Figure 1. Composite faces (bottom row) and shape grids (top row) depict the result of a multivariate regression testing associations between facial shape and aggressiveness. The image show the associated shape of low aggressiveness (left), the average (middle) and high aggressiveness (right). The image was taken with permission from Třebický et al. (2013).

2.4.2 Formidability, strength, dominance and masculinity. Not only the fWHR has been found to be related to perception of formidability, strength, dominance and masculinity. Adding to Carré & Mccormick (2008) findings of that aggressiveness can be accurately judged from fWHR, Třebický et al. (2013) identified several other facial features in MMA fighters that not only correlated with perceived aggressiveness but also with actual fighting success. In line with Carré & Mccormick (2008), faces that were judged as more aggressive had a broader bizygomatic width and shorter distance between eyes and mouth, indicating a higher fWHR in aggressive faces (see figure 1). However, other facial features such as more prominent eyebrows, horizontally narrower eyes, larger nose, broader chin, as well as lower fatty deposits, especially on the chin and cheeks also correlated with perceived aggressiveness (see figure 1). Furthermore, perceived aggressiveness was also found to correlate with actual fighting success ($r = .203$, $p = .01$). Together, these findings support that humans have evolved perceptual mechanisms for detecting not only fWHR but possibly also other facial cues related to formidability during physical male-male competition (Třebický et al., 2013).

A similar set of facial features as identified in Třebický et al. (2013) were also found in a study by Toscano, Schubert, and Sell (2014) who investigated the links between perceived strength, dominance and facial features. Adding to the finding of that people can accurately inference strength of others (Sell et al., 2009), Toscano et al. (2014) conducted two studies where participants rated not only how physically strong they thought individuals were but also how dominant. The first study used digitally generated faces and found that both strength and dominance were predicted by low brows, vertically narrow eyes, and wide noses and narrow mouths. Interestingly, the width of the face only predicted dominance but not strength. The second study used photos of real men and found that men with low brows, vertically narrow eyes and large chin were judged as both stronger and more dominant. However, most correlations in the second study were marginal, except for large chins predicting dominance ($p = .02$). Nonetheless, the significant correlations between facial features, perceived strength and dominance found by Toscano et al. (2014) overlap nicely with Třebický et al. (2013) and also fall in line with other studies on perceived dominance and strength (Dotsch & Todorov, 2012; Windhager et al., 2011;but also see Zebrowitz & Montepare, 1992).

Figure 2. Taken with permission from Windhager et al. (2011), the image visualise shape regressions showing facial shapes associated with certain characteristics.

Furthermore, Windhager et al. (2011) found that men that were judged more dominant and masculine by women had a wider and more prominent lower jaw as well as smaller eyes and nose compared to the averaged man. Men judged as more masculine and dominant also had smaller lips than the men judged as less dominant and less masculine. These features were also found to correlate with men rated as strong. However, men perceived as strong differed from men perceived as masculine and dominant in a few ways. Men perceived as strong had higher brows, larger eyes and less protruding chins (see figure 2) than men perceived as weak. This was taken to suggest that strength and dominance/masculinity judgments rely on similar cues but are independent of each other (Windhager et al., 2011), a suggestion fitting nicely with the findings by Toscano et al. (2014).

As the chin and jaw-line seem to be involved in perceptual judgments of aggressiveness, formidability, dominance, strength and masculinity, it could be expected that men with beards that change the outline if the chin and jaw would be judged differently from shaved men. Dixson, Barnaby, and Vasey (2012) found that women from different parts of the world rated the same men as less attractive, but of higher social status and older age when the men had beards compared to when they were clean shaved. When men displayed an aggressive facial expression, women rated men with beards as more aggressive in comparison with the same men displaying an aggressive facial expression while clean shaved. Given that beards are sexually dimorphic and seem to boost perception of aggressiveness in facial expressions, Dixson and Vasey (2012) suggest that beards might have evolved via intersexual selection for men to communicate status and aggression. Hence, beardedness seems to be a likely candidate to influence threat perception.

However, interestingly, Geniole and Mccormick (2015) found that fWHR predicted aggressiveness perceptions regardless of whether the men judged had beards or were clean-shaven. Predictions of masculinity ratings, on the other hand, were affected by beards. Interpreting their results, Geniole and Mccormick (2015) argued that while it is unknown whether beardedness acts as an accurate cue of aggression, there is evidence suggesting that fWHR does (see Geniole et al., 2015; Haselhuhn et al., 2015). Hence, humans might have evolved a preference for the upper facial regions related to fWHR and unaffected by beardedness when making judgments of aggressive tendencies in others (Geniole & Mccormick, 2015). Nonetheless, with relevance for the current study, beardedness still seems to influence judgments related to the lower face, as shown by its effect on masculinity ratings (Geniole & Mccormick, 2015).

2.4.3 Overgeneralisation, anger and fighting ability. There might be a link between perception of threatening traits in neutral faces and emotion recognition systems. The overgeneralisation hypothesis states that some neutral faces could be judged as more threatening than others because threatening neutral faces have features similar to the facial expressions of anger. That is, emotion recognition systems would overgeneralise angry-looking features to signal aggressive intent leading the observer to misattribute threatening dispositions, e.g., aggression, to the person (Knutson, 1996; Said et al., 2009). In support of the overgeneralisation hypothesis, Said et al. (2009) found that the features in neutral faces judged as threatening indeed positively correlated with features in angry faces and negatively correlated with surprised and happy facial features. Furthermore, neutral faces judged to have positive valence resembled happy faces and neutral faces with negative valence resembled fearful and disgusted faces. Hence, emotion recognition systems seem to overgeneralise features resembling emotional expression, consequently affecting how threatening a neutral face appears to be. However, the causal link might be reversed meaning that emotional expression could be perceived due to overgeneralisation by trait recognition systems (Said et al., 2009).

In line with the last mentioned suggestion, one explanation for why threatening neutral faces resemble angry faces is that the facial expression of anger could have evolved to display threatening characteristics in order to deter opponents (Sell, Cosmides, & Tooby, 2014). In support for of this view, Sell et al. (2014) found that features found in angry faces all independently influenced observers judgement of strength. Facial cues used to accurately infer others' strength are also used when assessing others' fighting ability (Sell et al., 2009). Hence, features in the angry facial expression might have evolved to cue strength and fighting ability. Important to note is that facial features in angry facial expression could also still express anger (aggressive intent) and not only strength and fighting ability. Contrary to the overgeneralisation hypothesis, it follows from Sell et al. (2014) that neutral faces with similar features as angry faces are not only found threatening because they resemble angry faces but also because such features signal physical strength. Hence, distinctions between perceptions of traits and perceptions of emotions appear to overlap.

2.4.4 Similar facial features occurring across studies. A study providing a link between the above-mention studies are Dotsch and Todorov (2012) who investigated what facial areas were used when people make judgements of dominance and trustworthiness of others. When participants made judgments of dominance and submissiveness, the facial areas used to make these judgments included the regions *around* the eyes, on the brows and the outline of the face, including the chin/jaw-line (see figure 3). Another construct associated with threat perception, namely trustworthiness, was shown to have similar, but slightly different diagnostic areas used for perceptual judgements. Areas used to make judgements about trustworthiness and untrustworthiness was *the eyes*, mouth and top of the head areas (see figure 3). As both dominance and trustworthiness are traits incorporated in threat perception (Dotsch & Todorov, 2012; but also see Geniole et al., 2015), it would be of relevance for the current study to control for the facial features used to perceive both dominance and trustworthiness. As it turns out, most of these areas, including the eyes and surrounding regions, brows and chin/jaw, are also often found in other studies were they are involved in perceptual judgements of aggressiveness, formidability (Dixson & Vasey, 2012; Třebický et al., 2013; Sell et al., 2009, 2014), strength (Sell et al., 2009, 2014; Toscano et al., 2014; Windhager et al., 2011), dominance (Toscano et al., 2014; Windhager et al., 2011) and masculinity (Geniole & Mccormick, 2015; Windhager et al., 2011).

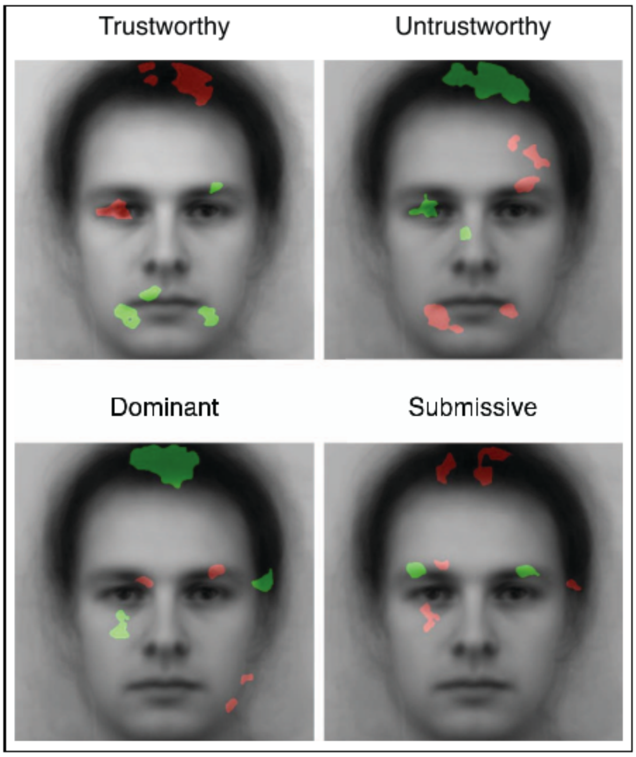

Figure 3. Depiction of the results from Dotsch and Todorov (2012), taken with permission. The areas highlighted were used to make corresponding perceptual judgments. Green areas represent where associations were made when the pixels of the area was displayed in high brightness; red areas represent where associations were made when the areas were dark.

Consequently, facial features that are strong candidates to positively influence threat perception, apart from fWHR, are: (1) the narrowness of the eyes (vertically and horizontally), (2) the protrusion, thickness and height of the brow, (3) the width of the chin and jaw, and (4) beardedness. However, perhaps the simplest way of explaining the links between fWHR and threat perception is that faces with high fWHR could share similar features with the facial expression of anger (Carré & Mccormick, 2008; Knutson, 1996; Said et al., 2009). Due to that angry faces have low brows (Langner et al., 2010) and previous studies on fWHR not controlling for brow height, could the entire link between fWHR and threat perception be due to overgeneralisation effects and/or possibly low brows signalling strength?

2.4.5 Questionable measurements. Almost all previous studies on fWHR and threat perception incorporate brow height in the fWHR by measuring upper facial height from the upper lip to the brow (e.g., Carré & Mccormick, 2008; but see Haselhuhn et al., 2015; and Geniole et al., 2015), confounding high fWHR with low brows. This is peculiar, as the foundational reasoning behind fWHR serving as a cue of aggression and threat (e.g., Carré & Mccormick, 2008; Geniole et al., 2015) relies on the finding by Weston et al. (2007) that the fWHR is sexually dimorphic, but Weston et al. (2007) did not incorporate brow height in the fWHR measure. Weston et al. (2007) measured the upper face from the nasion to the prosthion on skulls which leaves out the distance from the nasion to the brow ridge or to the brow in faces (see figure 4). With this in mind, the

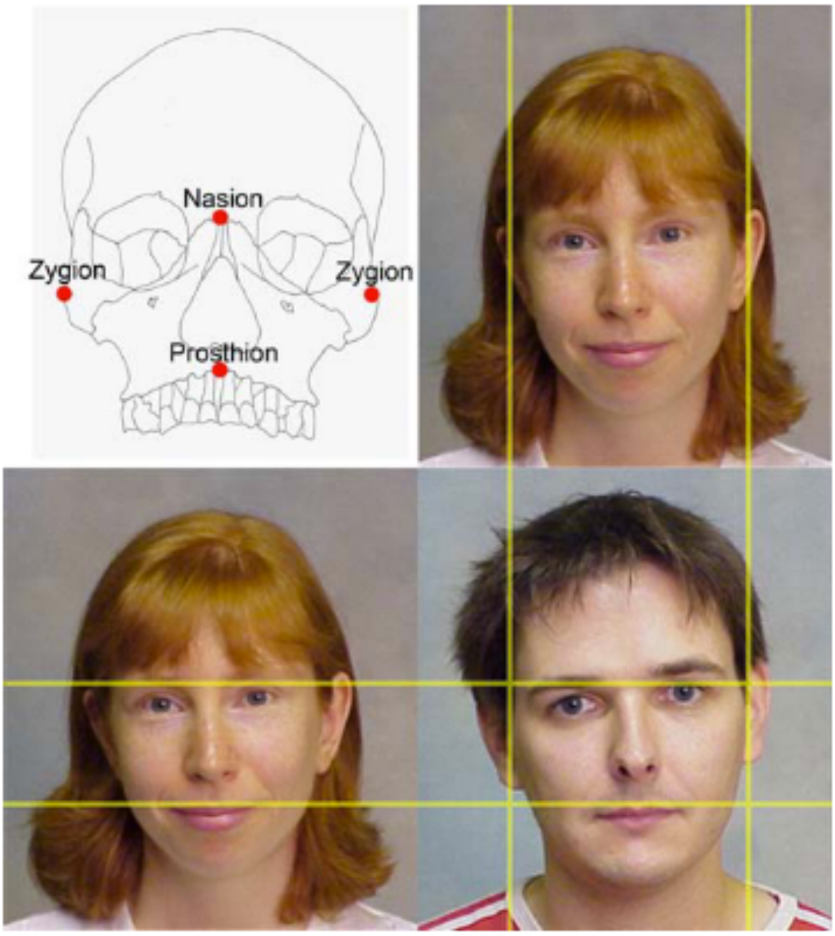

Figure 4. Image showing the reference points used by Weston et al. (2007) taken with permission. To illustrate the sexual dimorphism, the width and height of man's face are shown in relation to the woman's. Note the lines across the nasion crossing both the upper eyelid and the mid-brow on the man, illustrating the ambiguity of which feature to use as the upper limit of the facial height measure in photos.

effects of fWHR when including the distance from the nasion to the brow, as in previous studies, are not necessarily produced by the proposed sexually dimorphic nasion-to-prosthion distance. Instead, brow height could have produced most of the variance within these previous studies.

Surprisingly few studies to date seem to have left out brow height in their fWHR measure. Lefevre, Etchells, Howell, Clark, and Penton-Voak (2014) measured upper-facial-height from the upper lip to the upper eyelid and found significant correlations between fWHR and self-rated aggression in both men and women, linking fWHR with threat through aggression. However, Lefevre et al. (2014) also failed to replicate the sexually dimorphic link between fWHR and aggression (e.g., Carré & Mccormick, 2008; but see Haselhuhn et al., 2015) and threat (Geniole et al., 2015).

Another study (Gómez-Valdés et al., 2013) used a comprehensive method to estimate nasion-prosthion height, using markers from databases of cranial measures to model estimated locations of the prosthion and nasion that could be translated onto images. Gómez-Valdés et al. (2013) failed to link fWHR with aggression and found that the fWHR was not sexually dimorphic but instead more likely to be explained by the previous use of small sample sizes and that the previous links to aggression are more likely explained by socio-cultural factors. Interestingly, Kramer (2017) performed a meta-analysis on both skull measures from several international databases and also on faces used in previous studies and found no support for sexual dimorphism in fWHR when controlling for ethnicity and geographical origin. Hence, the common claim that fWHR is a sexually dimorphic feature might not be accurate, at least when not incorporating brow height into the fWHR measure.

In summary, while much support has been found for fWHR being an honest cue of threat (e.g., Geniole et al., 2015) there are strong claims against the fundamental base of the theory of fWHR being sexually dimorphic (e.g., Kramer, 2017). With other features such as brow height and jaw width also being known to influence threat perception, and possibly have confounded previous studies, an investigation of the validity of the fWHR as a measure influencing threat perception is warranted. Although, to do so, an objective measure of threat perception would be of substantial help as it could offer further validity to relationships found. In the following section, the possibility of doing precisely that will be explored.

2.5 The Late Positive Potential as an Objective Measure of Threat Perception

From an evolutionary perspective, psychological mechanisms, facial evaluation included, have been under selective pressure to use informative cues in the environment in order to produce adaptive regulation of biological processes (Tooby & Cosmides, 1990). Hence, cues of threat, e.g., a mean and angry looking conspecific, has been shaped by evolution to influence humans' emotional state, e.g., fear or anger, and in turn what action to take, e.g., flee or fight (Tooby & Cosmides, 1990). However, when asking a participant to evaluate a face by filling in a questionnaire, several methodological biases limit the accuracy of the participant's response (Podsakoff et al., 2003). It is possible that the participant initially could perceive an evolutionarily significant cue of threat but interpret the resulting emotional reaction as prejudice and rate the face in accordance with what they deem socially acceptable, rather than what they first thought. Hence, a method that objectively indexes the presence of evolutionarily significant cues would be desirable. A possible candidate to serve as such a measure is the ERP component called LPP (e.g., Schupp et al., 2000; Schupp, Öhman, et al., 2004; Van Strien, Eijlers, Franken, & Huijding, 2014; Weinberg & Hajcak, 2010).

When using EEG, the voltage across the scalp will be recorded, acting as a direct, although rough, measure of neural activity. Electrodes, placed on the scalp of the participant, pick up voltage changes caused by underlying neuronal activity, with minimal delay, allowing for accurate matching of fluctuations in the EEG data relative to a specific temporal event such as stimulus onset or a manual response. These event-related fluctuations in the EEG data are called ERPs (Cohen, 2017; Luck & Kappenman, 2011; Hajcak, Macnamara, & Olvet, 2010).

Specifically, what underlies the ERPs are postsynaptic potentials (PSPs) which is when a postsynaptic neuron has ions flow in or out of its cell membrane due to neurotransmitters binding to the postsynaptic neuron’s receptors (Cohen, 2017; Luck & Kappenman, 2011; Hajcak et al., 2010). This flow of ions makes the neuron a dipole with one end being positively charged and the other negatively charged. If PSPs occur in many neurons that are spatially aligned so that the negative and positive ends are oriented the same way, they can all be summarised into one large dipole called an equivalent current dipole. Partly due to the resistance of the tissue between the neuron and the electrode, thousands of spatially aligned neurons are required to create an equivalent current dipole great enough to be measured by EEG.

Fulfilling the above-stated criteria, cortical pyramidal cells are aligned perpendicular to the cortex and are believed to be the main source of ERPs (Cohen, 2017; Luck & Kappenman, 2011). If the neurons in a given region are not oriented the same way, e.g., as in the basal ganglia, most of their individual dipoles will cancel each other out. Interneurons also seem to have little to no effect on scalp-recorded ERPs. Hence, a very limited measure of brain activity is captured by the ERPs which typically reflect simultaneous PSPs in large populations of similarly oriented cortical pyramidal neurons (Cohen, 2017; Luck & Kappenman, 2011).

As PSPs occur throughout the brain at all times due to the brain being constantly active, the EEG data constantly fluctuates which makes it problematic to separate fluctuations produced by ERPs from fluctuations not related to ERPs (Luck, 2014). Fluctuations in the EEG data not related to ERPs often oscillate between negative and positive voltage changes, reflecting feedback loops in the brain. By recording EEG data over several trials, presenting stimuli with irregular intervals, and then averaging the EEG-data across the trials cancels the effect by the oscillations while leaving the effect by ERPs visible. This makes ERPs more detectable when depicted an in an average ERP waveform.

ERP waveforms are depictions of scalp-recorded voltage (EEG data) over time, relative to a temporal event, depicted as a continuous waveform over a pre-specified time window. The waveform can include positive and negative peaks reflecting each electrodes local voltage maxima with variation in amplitude and duration (Luck & Kappenman, 2011). Specific ERP peaks are associated with, but not equal to, ERP components which in turn can be defined as ”*a scalp-recorded voltage change that reflects a specific neural or psychological process*” (Luck & Kappenman, 2011, pp. 4). The voltage change produced by PSPs of a specific neural process, that is, an ERP component, will linger for tens to hundreds of millisecond, making each specific peak represent the sum of voltage change from temporally overlapping ERP components (Luck & Kappenman, 2011). Furthermore, voltages distribution across the scalp also spreads laterally when passing through the skull, due to the skulls relatively high resistance. Hence, it becomes impossible to calculate the neural source of the voltage change.

However, ERP components are often related to specific scalp distributions of activity, which can be calculated using, e.g., independent component analysis (ICA) which uses the data’s statistical properties to identify likely independent components’ scalp distribution. Taken together, the latency of an associated peak, the peak’s polarity, the scalp distribution and what stimuli or manipulation produce it and in what context, can be used as an *operational definition* of an ERP component (Luck & Kappenman, 2011).

Regarding the LPP, it has been described as a measure indicating intensity of emotional arousal to motivationally salient stimuli of evolutionary significance and possibly sustained attentional resources (Hajcak et al., 2010; Lang, Bradley, & Cuthbert, 1997; Schupp, Cuthbert, et al., 2004; Schupp, Junghöfer, Weike, & Hamm, 2004; Schupp, Öhman et al., 2004; Weinberg & Hajcak, 2010). The LPP has been found to increase after observation of both pleasant and unpleasant words, faces and pictures when compared to neutral versions of the same stimuli (MacNamara, Foti, & Hajcak, 2009, 2010; Schupp, Cuthbert et al., 2004; Schupp, Junghöfer et al., 2004; Schupp, Öhman et al., 2004; Weinberg & Hajcak, 2010). Hajcak et al. (2010) suggest that LPP respond to stimuli of subjective value, linking LPP to intrinsic motivation. Typically, LPP starts around 300 ms after stimuli onset and has been found to extend the full length of stimuli presentations lasting up to as long as 1500 ms (Cuthbert, Schupp, Bradley, Birbaumer, & Lang, 2000; Hajcak & Olvet, 2008) and even after stimulus offset (Hajcak & Olvet, 2008).

Furthermore, the scalp topography of LPP typically spans the centroparietal midline, starting at the parietal midline moving anteriorly until reaching a more central distribution (Foti, Hajcak, & Dien, 2009; MacNamara et al., 2009) The source of the voltage change has been linked to underlying neural activity in the occipital and posterior parietal cortex (Hajcak et al., 2010; Keil et al., 2002). Furthermore, detection of novelty or presentation of unexpected stimuli does not seem to affect LPP (Hajcak et al., 2010), neither does the LPP habituate over the course of repeated emotional stimuli presentation (Codispoti, Ferrari, & Bradley, 2006, 2007; Olofsson & Polich, 2007). This is interesting as other measures of arousing reactions, like skin conductance response, habituate to arousing stimuli which suggest that the LPP might reflect early forms of categorisation processing, instead of reflecting emotional reaction (Codispoti et al., 2006, 2007). This makes the LPP promising as an index of perception of evolutionarily relevant cues of threat.

Based on findings from principal component analysis (PCA), an analytical tool using statistical properties of EEG data to identify ERP components and their scalp distributions, MacNamara et al. (2009) suggests that the LPP might be best described as the sum of several overlapping components related to emotional modulation. Others have suggested and found support for that the early ERP components, occurring around < 300 ms after stimulus onset reflects attentional capture while the later LPP component (>300 ms) reflect more flexible and sustained processes (Foti et al., 2009; Weinberg & Hajcak, 2010). Furthermore, highly emotionally salient stimuli, e.g., erotica and mutilation-objects, have been found to produce the larger LPPs than less emotionally salient stimuli, e.g., affiliative-objects and disgust-objects (Weinberg & Hajcak, 2010). Moreover, differences between early (400-1000 ms) and late (1000-1500 ms) latencies of the LPP was found, showing that erotica produced exceptionally high magnitudes in the earlier latency window while being comparable to other stimuli in the later window (Weinberg & Hajcak, 2010).

Relevant to the current study, these results seem to translate to when threatening faces are used as stimuli (Eldblom, 2018; Schupp, Öhman et al., 2004).

Schupp, Öhman, et al. (2004) found that LPP amplitudes were greater when participants observed faces expressing anger compared with when participants observed neutral or friendly facial expressions. Interpreting their findings, Schupp, Öhman, et al. (2004) argued that they supported that evolutionary pressures shaped perceptual systems to direct attention towards stimuli of evolutionary significance. Further support for this view has been derived from that the LPP magnitudes were greater when participants observed both pleasant and unpleasant evolutionary significant pictures of (e.g., erotica and mutilated limbs) when compared with pictures that had an equally strong emotional valance but lacked imminent or direct evolutionary significance for the individual (e.g., contamination) (Schupp, Cuthbert, et al., 2004; Schupp, Junghöfer, et al., 2004; Schupp, Öhman, et al., 2004; Weinberg & Hajcak, 2010). Furthermore, dangerous animals, such as snakes, have been found to elicit greater LPP amplitudes in comparison with less dangerous animals, such as birds (Van Strien et al., 2014). Taken together, these above-mentioned studies build a compelling case for LPP indexing cues of evolutionary significance as seen by its relation to emotionally salient stimuli, including threatening stimuli.

Apart from the LPP, other ERP-components could be influenced by threat perception as well. The first positive peak after stimulus onset in figure 8 reflects the N170/vertex positive potential (VPP), a component associated with face stimuli but have been found to be modulated by emotional expression with angry faces eliciting higher amplitudes (Batty & Taylor, 2003). The second negative peak, ca 300 ms after stimulus onset, in figure 8 is the early posterior negativity (EPN) which have been linked to emotional stimuli with selective attention properties (Luck & Kappenman, 2011). The second positive peak after stimulus presentation seen in figure 8 is the P300, a component associated with higher magnitudes for target stimuli in research paradigms demanding response to target stimuli, infrequently presented stimuli (Duncan-Johnson & Donchin, 1977) and emotionally salient stimuli respectively (Schupp, Öhman et al., 2004). However, regarding the current study, only the LPP will be investigated as an index of threat due to its repeatedly documented connection to evolutionary relevant and motivationally salient stimuli (e.g., Cuthbert et al., 2000; Schupp, Cuthbert et al., 2004; Schupp, Junghöfer et al., 2004; Schupp, Öhman et al., 2004; Van Strien et al., 2014; Weinberg & Hajcak, 2010).

The idea to connect LPP with the fWHR was taken from Eldblom (2018) who found that when participants were cued to perceive faces in a threatening context, the LPP was greater for faces with high compared to low fWHR. However, as fWHR varied without controlling for variance in other facial features related to threat, the link between fWHR and LPP found in Eldblom (2018) might be confounded by variation in these other threat-related facial features. Furthermore, as LPP does not exclusively respond to threatening stimuli but also seems to respond to other stimuli of evolutionary significance, LPP might be able to indicate support for what facial features humans are biased towards when perceiving and evaluating faces.

2.6 Rationale For The Present Study

As specified earlier, there is substantial support for that the fWHR functions as an honest cue for threat but there are also other facial features that could influence how threatening a face is evaluated to be. Therefore, there is a need to test whether the fWHR truly is an honest cue of threat or whether the influences of other facial features have confounded the fWHR measure in previous studies. In order to do so without sacrificing too much statistical power, the current study will only focus on three facial features that can be argued to have a strong chance of influencing threat perception of a face. These three features include differences in: (a) cheekbone width, which will represent differences in the fWHR measure; (b) eyebrow height, changing independently from the fWHR (using the prosthion and naison, and not the brow or upper eye-lid as the boundary for the upper facial height); and (c) the width of the jaw, also changing independently from the fWHR. Consequently, the current study will investigate whether the height of the eyebrows and the width of the jaw confound of the expected difference in threat perception between wide cheekbones (high fWHR) and slim cheekbones (low fWHR).

Subjective ratings can be problematic as the participant could modulate their response due to experimenter effects, cultural or social factors, an objective measure indexing threat perception would be useful. As previously described, there are reasons to believe that the LPP could be used to serve as such a measure. In order to further explore the possibility of using LPP to objectively index threat perception, the current study aims to compare facial ratings with measures of the LPP of participants watching faces that differ in cheekbone width, eyebrow height and jaw width. The primary hypothesis is that faces having either wide cheekbones, and/or low brows, and/or a wide jaw will induce a higher average LPP and be rated as more threatening than faces having low brows, and/or slim cheekbones, and/or slim jaw. The secondary hypothesis is that eyebrow height will have the greatest effect on both facial ratings and average LPP magnitude.

3. Method

3.1 Participants

Participants (N = 30, females = 11, M_{age} = 24 years, SD_{age} = 2.9) consisted of undergraduate and graduate students of the University of Skövde, including both Swedish and international students. The exclusion criteria required participants to: (a) be between 18-40 years old; (b) be right-handed; (c) have good sight and be able to clearly see 2 m ahead, including with glasses or contact lenses; (d) not suffer from epilepsy; (e) not have any current ongoing psychiatric or neurological illness or diagnosis; (f) not be colour-blind; (g) speak English; and (h) not be dyslexic.

3.2 Sampling Procedure

The sampling of participants was carried out through scripted invitations, either posted on social media or e-mailed to students, with the incentive that participants would be shown their own brainwaves. The same incentive was also given to students approached in person at the University of Skövde campus. Participants received no monetary compensation. Sampling was carried out over 27 consecutive days. Prior to participation, all participants received information about the study in accordance with the Declaration of Helsinki and consented to participation.

3.3 Stimuli

The primary stimuli consisted of eight distinct computer-generated neutral male faces which underlying facial bone structure had been adjusted in FaceGen (FaceGen Modeller Core 3.18, Singular Inversions, 2018). Together, the primary stimuli faces exhibited all combinations of the variables (a) low vs. high brows, (b) wide vs. slim cheekbones, and (c) wide vs. slim jaw. The low brows, wide cheekbones and wide jaw represented the threatening features (see figure 5) and were modelled after faces found to be perceived as threatening, strong, dominant, masculine or aggressive (Carré & Mccormick, 2008; Třebický et al., 2013; Windhager et al., 2011). The low brows, slim cheekbones and thin jaw represented the less-threatening features (see figure 5) and

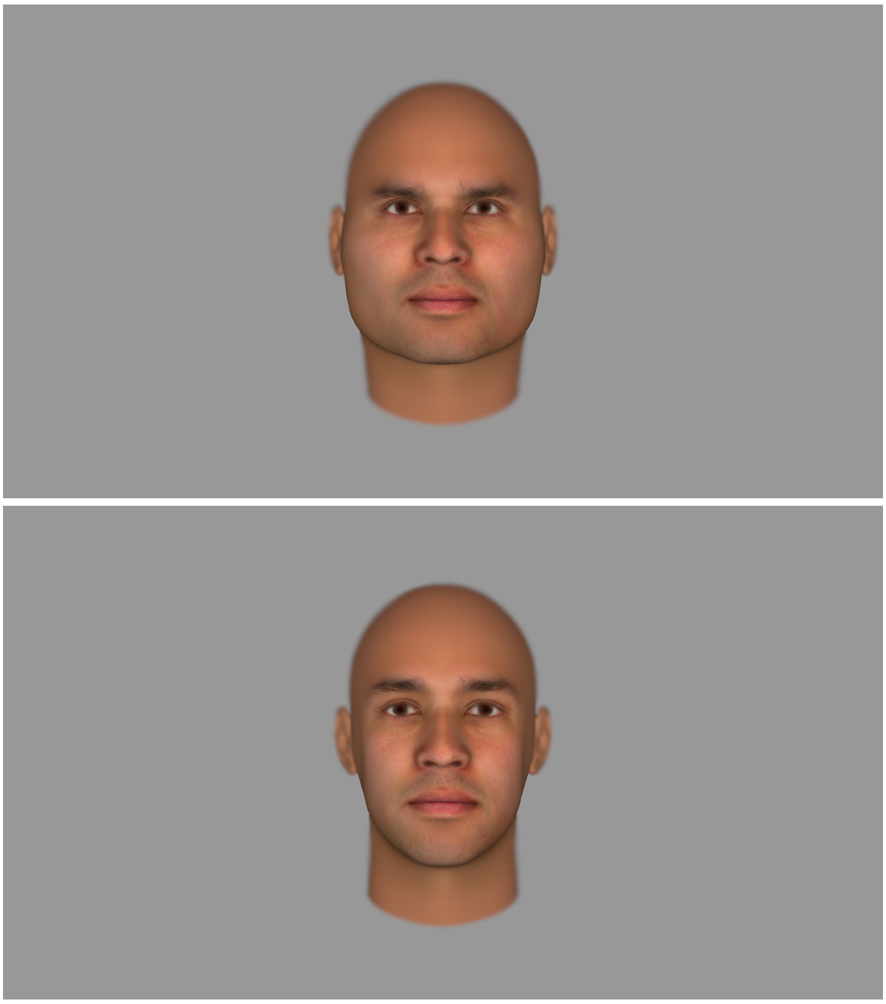

Figure 5: Two faces from the primary stimuli. The top face exhibits all threatening features. The bottom face exhibits all less-threatening features.

were modelled after faces found to be perceived as either less threatening, weak, submissive, feminine or less aggressive (Carré & Mccormick, 2008; Třebický et al., 2013; Windhager et al., 2011). Additionally, replicas were made of each face of the primary stimuli but these additional eight faces had closed eyes and were only used as secondary stimuli for the attention task,. For the exact settings used in FaceGen for each variable and face see appendix A. After being created in FaceGen, all faces were blurred around the ears, scalp and neck to prevent attention being drawn to these otherwise more unnatural-looking areas (see figure 5).

Using an E-prime 2.0 software (Psychology Software Tools, 2019) model, the primary stimuli were 8x8 pseudo-randomised over 12 blocks each consisting of 64 trials with no identical faces being presented two times in a row. The attention-task stimuli were randomised into every block to occur one-eighth of the time. Each trial had a stimulus onset asynchrony (SOA) of 2200-2400 ms and followed the model seen in figure 6. The background was grey and its illuminance matched the faces', preventing illuminance differences to act as a confound. Otherwise, illuminance differences could have occurred due to the varying sizes of the faces which could possibly confound the effects of the stimuli itself.

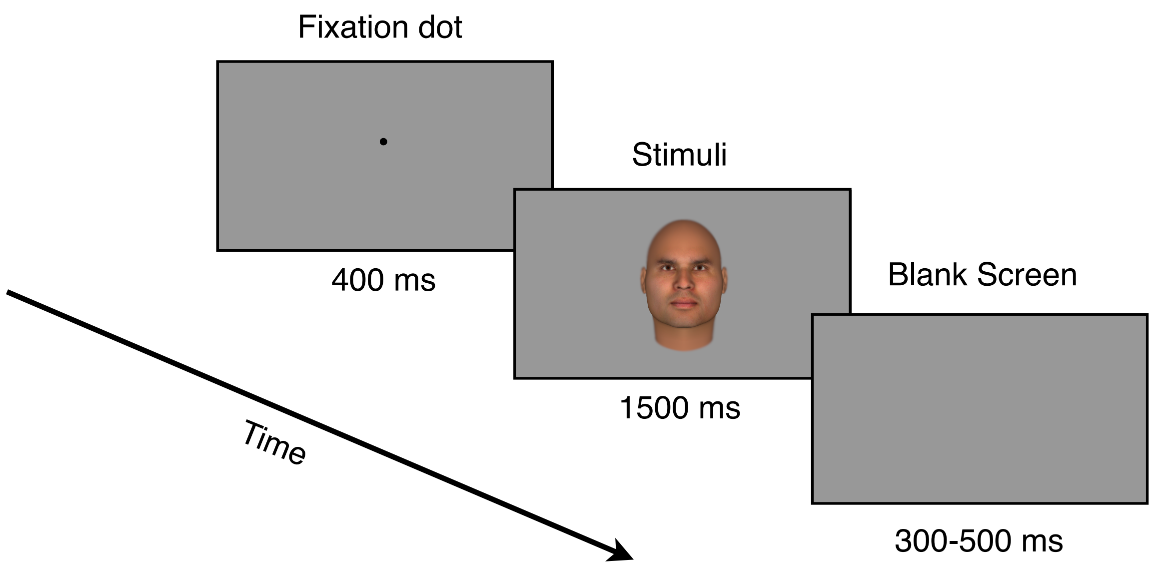

Figure 6. Representation of the model used for stimuli presentation showing one trial. The fixation dot was positioned to occur in the same area as in-between the eyes of the faces.

3.4 Facial Ratings and Questionnaires

A questionnaire assessing subjective threat perception was used, making participants rate the primary stimuli on a Likert scale ranging from 1 (*Least threatening)* to 9 (*Most threatening*). Participants also filled out three other questionnaires on personality, anxiety and aggression. However, as the data collected by these three additional questionnaires were for other studies and are considered elsewhere, the current study will not consider them further.

3.5 Procedure

Upon arrival, the participation information was read by the participants followed by they giving their informed consent to participation. Afterwards, participants were led into a separate dim-lit room where the temperature was kept stable at approximately 24-25º Celsius. The chair, in which the participants were seated, was positioned so that the participants' eyes were approximately 110 cm in front of a 24", 1920×1080p resolution, 60Hz frame rate, computer monitor (HP Compaq LA2306x) on which the questionnaires and stimuli later were presented. After the head measurements were taken, the head cap and electrodes were attached. While the electrodes were gelled, the participant started filling out the questionnaires using the number keys on a keyboard. When the questionnaires and gelling were done, the keyboard was switched to a hand-held controller followed by on-screen instructions telling the participant to press a button when seeing a face with closed eyes, that is, the secondary stimuli. Pressing the button at the right stimuli constituted the attention task. Button presses were recorded to control for the participants' alertness. Upbeat music, previously chosen by the participants for the experiment, was playing during both the questionnaires and the presentation of the stimuli in order to keep participants alert. The participants chose the music volume before the experiment started. After every block of stimuli, there was a short break for participants to rest while on-screen instructions said to press the button when ready to continue. Halfway through all blocks, after the 6th block, there was a longer break where participants were invited to leave their chair and receive a glass of water and a cookie if they wanted to. After finishing the final blocks, participants were debriefed.

3.6 EEG Recording Setup

A total of 17 active Ag/AgCl electrodes (g.LADYbird electrodes, manufactured by g.tec) were used whereof 13 were mounted on a stretchable cap (g.GAMMAcap[3]) and positioned in accordance with the following international 10/20 Placement System positions: AF4, AF3, Fz, FC4, FC3, Cz, CPz, CP2, CP1, Pz, P6, P5, and Oz. All other electrodes were attached with adhesive tape. The positions were chosen with regards to the LPP scalp distribution as well as to have enough spread to enable independent component analysis (ICA). The active electrode impedances were transformed by the system to output impedances of about 1kOhm. Electrodes were online referenced to the right mastoid (RM), and FPz served as ground. The data was acquired in MATLAB R2015a (version 8.5.1.281278; The MathWorks, inc., 2019) with the g.USBamp (g.tec) amplifier. It was sampled at 256 Hz on a digital signal processor within the amplifier, and filtered online with an eighth-order Butterworth lowpass filter with a half-power (-3dB) cutoff at 60 Hz. Two electrodes were attached to the left and right mastoids for later offline re-referencing. Ocular movements were captured by attaching electrodes at the external canthi and suborbit of the right eye.

3.6.1 Processing of EEG-data. Offline analysis was performed using the toolboxes EEGLAB (version 13.6.5b; Delorme, & Makeig, 2004) and ERPLAB (version 7.0; Lopez-Calderon, & Luck, 2014) in MATLAB. Continuous EEG data were re-referenced to the average of the mastoids and filtered offline with a 180th-order stopband notch filter at 50 Hz (to remove line noise). As a preprocessing step for removing artifacts with Independent Component Analysis (ICA), the data were filtered with a second-order Butterworth bandpass filter with a half-power (-3dB) cutoff at 1 and 30 Hz. The EEG data were then segmented into epochs of 1900 ms, with 400 ms pre-stimulus baseline and 1500 ms post-stimulus. Thereafter, epochs exceeding three standard deviations above joint electrode probability activity limits were rejected. ICA was then run, and Multiple Artifact Rejection Algorithm (MARA; 2013) was used to automatically identify and remove ICA components reflecting artifacts (Winkler, Haufe, & Tangermann, 2011). Following this, the ICA weights were transferred back to the pre-processed, unepoched data that had only been subjected to the notch filter, and the relevant MARA-detected components removed from this data. Subsequently, the data were filtered with a second-order Butterworth highpass filter with a half-power (-3dB) cutoff at 0.1 and data was, as before, segmented into epochs of 1900 ms, with 400 ms pre-stimulus baseline and 1500 ms post-stimulus. Step-wise artifact rejection was performed in ERPLAB 7.0 (epochs containing steplike activity greater than 100 µV in a moving window of 200 ms with a step size of 20 ms were rejected). Subsequently, epochs were averaged for each participant and each experimental condition, and low pass filtered at 30 Hz to ease visual inspection.

Due to prior research consistently finding the LPP to be most pronounced over central-parietal sites (e.g., Foti et al., 2009; MacNamara et al., 2009; Weinberg, & Hajcak, 2010), the LPP was quantified across a cluster of central-parietal electrodes (Cz, CP1, CP2, CPz, Pz) as a function of the condition spanning an early (400-800 ms) a late (800-1,500 ms) time window after stimulus onset. The time windows were chosen with regard to previous research (Eldblom, 2018; Schupp, Junghöfer, et al., 2004; Schupp, Öhman et al., 2004; Weinberg, & Hajcak, 2010).

4. Analysis

A repeated measures ANOVA was performed on the facial rating scores and the ERP-measures respectively, including three factors with two levels: brow height (low vs. high), cheekbone width (slim vs. wide) and jaw width (slim vs. wide). Effect sizes from the repeated measures were reported in partial eta squared (η_p^2). The alpha-level was set $\alpha = .05$ (two-tailed) for all analyses. One participant's data were excluded from the facial ratings due to misunderstanding the task. Nine participants were excluded from the EEG-measures analyses due to bad data quality on the basis of having over 20% of their trials rejected during data processing (see table 1 in appendix B). For the facial ratings, a second analysis excluded outliers either 3 inter quartile ranges (IQR) bellow the 1st quartile, or 3 IQR over the 3rd quartile. In line with recommendations from Luck (2014), no outliers in the EEG-data were excluded from the analysis. All analyses were performed in SPSS Statistics (version 25).

5. Results

5.1 Facial Ratings

The repeated measures ANOVA revealed a significant main effect for brow height $F(1{,}28) = 64.823$, $p < .001$, $\eta_p^2 = .698$ but not for cheekbone width $F(1{,}29) = .867$, $p = .36$, $\eta_p^2 = .030$ nor jaw width $F(1{,}29) = .020$, $p = .89$, $\eta_p^2 = .001$. No interaction effect was significant. See figure 7 for overview. Removing an outlier over 3 IQR over the 3rd quartile strengthened the previously observed significant main effects for brow height $F(1{,}27) = 64.556$, $p < .001$, $\eta_p^2 = .705$ but still not for cheekbone width $F(1{,}27) = .867$, $p = .36$, $\eta_p^2 = .031$ nor for jaw width $F(1{,}27) = .002$, $p = .96$, $\eta_p^2 < .001$. No interaction effect was significant.

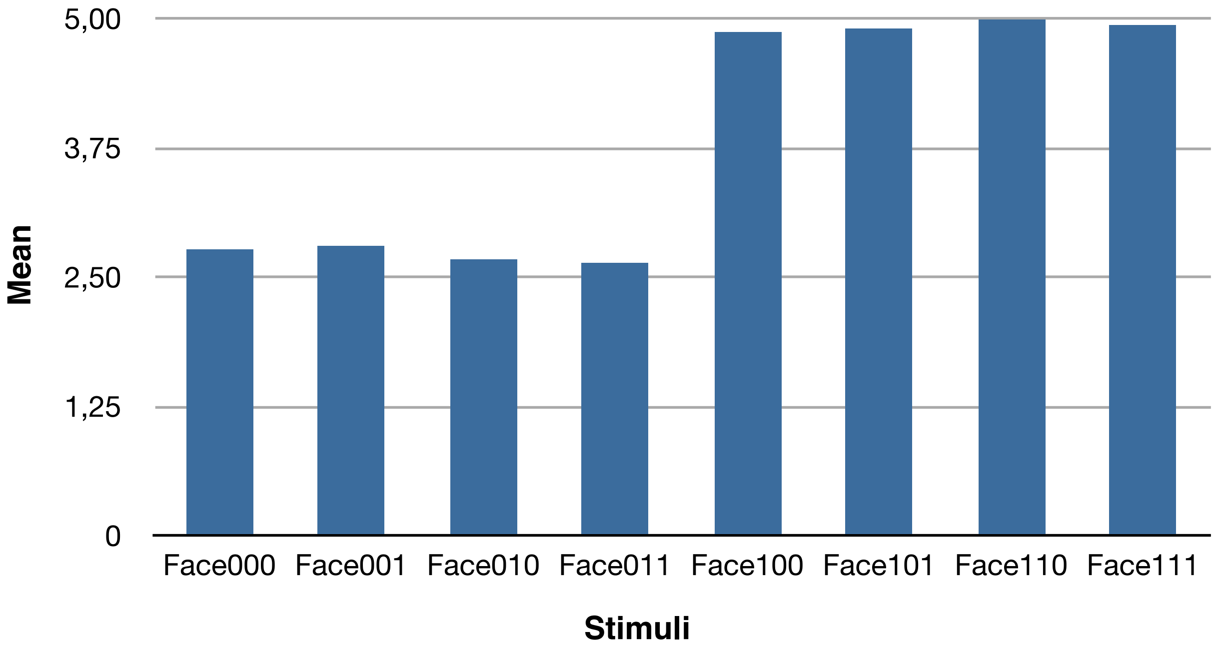

Figure 7. Bar graph illustrating mean scores of the facial rating. The numbers in the stimuli names describe the faces' variable settings. The order of the numbers indicate the variable: brows first, cheekbones second and jawline third. The numbers themselves indicate the variable level, 1 equals threatening-feature and 0 less-threatening-feature (see figure 5).

5.2 EEG-Measures

5.2.1 Entire LPP (400-1500 ms). The repeated measures ANOVA revealed no significant main effects for neither brow height $F(1,20) = 1.979$, $p = .18$, $\eta_p^2 = .090$, cheekbone width $F(1,20) = .120$, $p = .73$, $\eta_p^2 = .006$ or jaw width $F(1,20) = .014$, $p = .60$, $\eta_p^2 = .014$. No interaction effects was found.

5.2.2 Early LPP (400-800 ms) vs. late LPP (800-1500 ms). When time was incorporated as an factor (early LPP, 400-800 ms window; and late LPP, 800-1500 ms window) the repeated measures ANOVA revealed a significant main effect for brow height $F(1,20) = 5.403$, $p = .03$, $\eta_p^2 = .213$ but not for cheekbone width $F(1,20) = .148$, $p = .70$, $\eta_p^2 = .007$ or jaw width $F(1,20) = .457$, $p = .51$, $\eta_p^2 = .022$. However, brow height had a significant interaction with cheekbone width $F(1,20) = 21.425$, $p < .001$, $\eta_p^2 = .161$ with slim cheekbones and low brows resulting in larger LPP magnitude. Brow height also had a significant interaction with jaw width $F(1,20) = 6.045$, $p = .02$, $\eta_p^2 = .232$, with low brow and slim jaw generating larger LPP magnitude.

5.2.3 Post hoc findings. Upon inspection of the plot of the grand average ERP (figure 8 and 9) and difference waves (figure 10), only the LPP of faces with threatening brows (low position) and LPP of faces with less-threatening brows (high position) is clearly visually distinct from one another between the 400-800 ms window. A paired samples t-test revealed that LPP magnitude elicited by low brows differed significantly from high brows in the early LPP (400-800 ms) $t_{20} = -2.485$, $p = .02$, $d = .542$, $M = -.592$, $SD = 1.092$, but not in the late LPP (800-1500 ms) $t_{20} = -.711$, $p = .49$, $d = .155$, $M = -.185$, $SD = 1.190$.

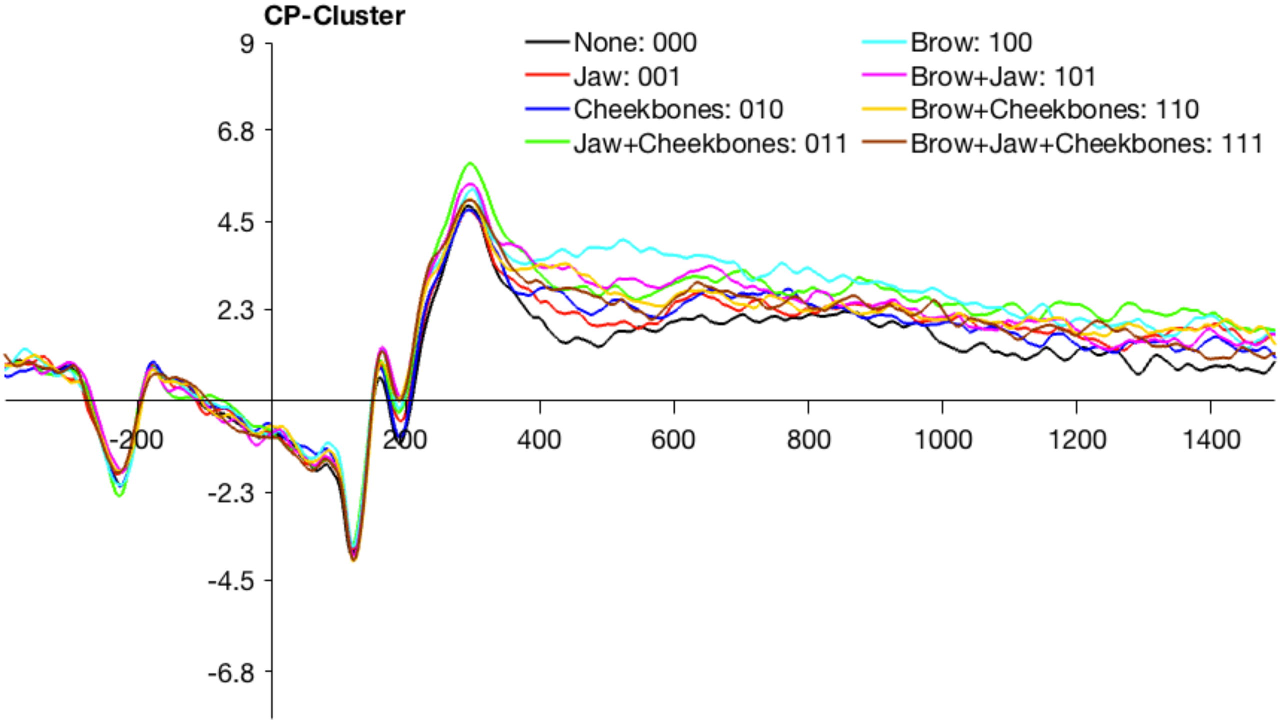

Figure 8. Plot of the grand average LPP for each face stimuli. To ease overview, the descriptions of the stimuli describes what variable/s were set to the threatening-feature level as well as the corresponding binomial name used in figure 7.

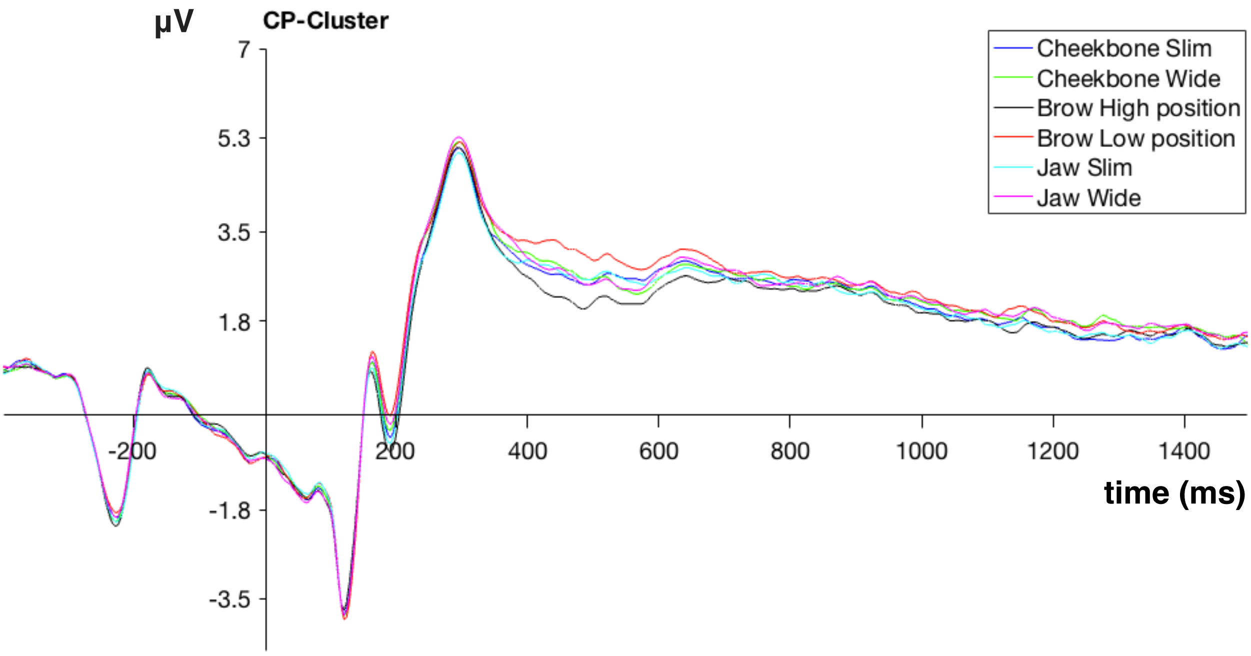

Figure 9. Plot of the grand average ERPs for each level of each variable.

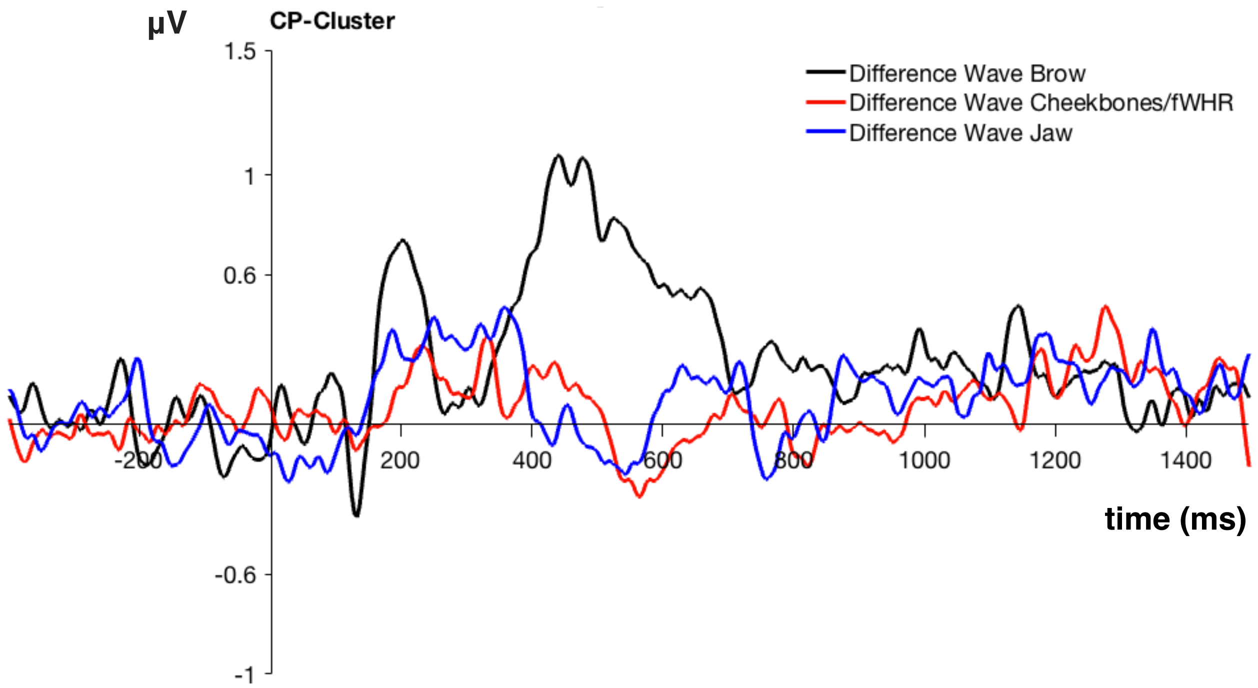

Figure 10. Plot of the difference wave for: brow (low position - high position), cheekbones/fWHR (wide - slim) and jaw (wide - slim).

6. Discussion

Little to no support was found for the primary hypothesis stating that all faces with threatening features (low brows, wide cheekbones and wide jaw) would all respectively affect facial threat ratings and LPP magnitude. However, only brow height had a significant effect on facial ratings and LLP magnitude. Nonetheless, this is in line with the secondary hypothesis of brow height producing the greatest effects. However, as brow height only had a significant effect on LPP magnitude in the early LPP (400-800 ms), possible explanations for the LPP attenuation after this time window should be addressed before any inferences are made.

Concerning the LPP, a time window from 400-800 ms is well within the boundaries of previous research related to threatening content, e.g., 350-750 ms (Schupp et al., 2000), 400-600 ms Schupp, Öhman, et al. (2004) has been used. On the other hand, others have used the window 400-600 ms to measure how threat-related content effects the P300 (Schupp et al., 2007), an ERP component associated with the positive peak approximately 300 ms after stimulus onset. Exactly where to separate the P300 and the LPP is unclear and without no consensus in the literature as components can easily overlap (Hajcak et al., 2010; Luck, & Kappenman, 2011). However, in a study using principal component analysis (PCA), a method that can separate distinct neural responses with temporal overlap, it was suggested that the LPP is the sum of several overlapping positive components, the P300 included (Foti et al., 2009, 2009). Hence, it seems preferable to not view the LPP, and especially the early LPP, as a sole, independent, component generating sustained positivity.

However, the sustained positivity is distinctive for the LPP and has been suggested to be due to motivational engagement and sustained attention (Foti et al., 2009, 2009; Lang et al., 1997) while the LPP's magnitude is associated with how emotionally arousing (e.g., Cuthbert et al., 2000; Schupp, Cuthbert, et al., 2004; Weinberg & Hajcak, 2010) and evolutionary significant the stimuli is (e.g., Schupp, Cuthbert, et al., 2004; Schupp, Öhman, et al., 2004; Weinberg & Hajcak, 2010). Given that the stimuli used in the current study were of neutral faces, and that the low and high levels for brow did not differ greatly when compared to angry and neutral faces in previous studies (e.g., Schupp, Öhman, et al., 2004), a less sustained LPP of lower magnitude would be expected in the current study.

6.1 Cheekbone Width, Brow Height and fWHR

Concerning the primary hypothesis, the current study failed to replicate previous findings (e.g., Geniole et al., 2015) of higher rated perceived threat in faces with relatively wide cheekbones, meaning high fWHR. Even though most of the other studies linking high fWHR to threat perception used real photos or composites of real photos, it is unlikely that the present study's results would be explained only by that the stimuli were computer-generated. Another study using the same software (FaceGen) as the current study to generate faces found strong support for fWHR being a perceptual cue of threat (Carré, Morrissey, Mondloch & McCormick). Furthermore, FaceGen has been successfully used in several studies (e.g., Oosterhof & Todorov, 2008; Sell et al., 2014; Toscano et al., 2014) linking facial features to perceptual judgments of threat, strength and dominance. Rather than being due to using computer-generated faces, a more probable explanation for why the current study did not replicate previous fWHR findings can be found in the facial measures used in previous studies.

Most studies on fWHR as incorporated brow height in their measure of the upper facial height. With this in mind, and that the theory of fWHR being a cue of threat that evolved from male-male competition pressures have been undermined by lacking evidence of sexual dimorphism, it is reasonable to believe that part of the variance seen in these studies were actually driven by brow height. That said, some evidence still points towards fWHR affecting judgement. Stirrat and Perrett (2010) used the height from the upper lip to the eyelid to mark the upper facial height. Despite not incorporating brow height they found that faces with lower fWHR were judged as more trustworthy than faces with higher fWHR. Interestingly, Stirrat and Perrett (2010) presented the pictures simultaneously in pairs to their participants who then had to choose which face they thought seemed most trustworthy. However, making participants compare faces might have enhanced the effects of the manipulated fWHR as the faces are seen in relation to each other. An interesting future study would be to use the same fWHR manipulation but to present the images one by one, as in the present study.

Another possibility would be that when participants can compare two pictures and are not time restricted in making their judgements as in Stirrat and Perrett (2010), then a greater amount of subtle features could be incorporated in the processing and influence the perceptual judgments. If this was the case, it could potentially explain why the current study did not find evidence of cheekbone or jaw width affecting threat perception. However, only 39 ms are needed for people to make consistent judgements of threatening personality in neutral faces (Bar, Neta, & Linz, 2006). Interestingly, as judgments after 39 ms are consistent in pictures with low spatial frequencies, the early processing of threat in faces seem to rely solely on rough outlines of features such as the height and angle of the brows (Bar et al., 2006). Hence, while brow height might influence threat judgements after as little as 39 ms, other features might not be as sensitive (Bar et al., 2006).

Moreover, judgments of other character traits were not as consistent in the early processing of faces in Bar et al. (2006) but consistent judgements of traits such as trustworthiness (Oosterhof & Todorov, 2008; Stirrat & Perrett, 2010) and strength (Sell et al., 2009) has been reported when participants could look at the pictures for as long as they wished. It is possible, then, that judgements of trustworthiness and strength might relate to features such as cheekbone and jaw width but that incorporating these features into perceptual judgments require more processing resources. In regards to theories of how evolutionary pressures form perceptual biases towards stimuli threatening to survival (e.g., Tooby & Cosmides, 1990; Sell et al., 2009), it could be argued that while cues of others' motivational states towards you (e.g., an angry person staring at you signalling aggressive intent) constitutes a more immediate threat than another person's strength or whether they have a trustworthy personality. The same argument would also support overgeneralisation effects having greater influence on threat perception than cues of strength.

6.2 Jaw Width and Threat Perception

Regardless, the current study's results lend no support to that faces with greater jaw width, due to associations between jaw width and strength (Windhager et al., 2011), would be perceived and judged as more threatening than a face with slimmer jaw width. Strong support has been found for people having the ability to accurately infer the strength of men's faces (Sell et al., 2009; Windhager et al., 2011) and that inference of strength seem highly relevant from an evolutionary perspective of male-male competition (Sell et al., 2009, 2012). Therefore it is surprising that no effect for jaw width was found in the current study.

The simplest explanation would for this be that jaw width is not an evolved cue of strength but that other facial features instead influence judgements of strength. Interestingly, Toscano et al.,

(2014) found that both low brows, larger chin and the how vertically narrow the eyes were positively correlated with perceptions of strength and dominance. Perhaps the narrowness of the eyes and the width of the chin influence strength and dominance perception to a greater degree than jaw width. Upon inspection of the stimuli used in studies were actual fighting ability correlated with subjects ratings of formidability (Třebický et al., 2013; Zilioli et al., 2015) and aggression (Třebický et al., 2013), it is clearly evident that facial stimuli judged as more formidable also have vertically narrower eyes and lower brows. Investigating judgments on strength in neutral faces while controlling for the narrowness of the eye and possibly chin size would be an interesting future study.

6.3 Interaction Effects

Concerning the significant interaction effect showing that low brow position and slim cheekbone width generated higher LPP, the fWHR's connection to threat is again undermined. What the interaction could be due to is unclear. Possibly, the face with slim cheekbones looked more athletic compared to the face with wider cheekbones. However, whether looking athletic would have an effect is highly speculative. Regarding the interaction with low brows with slim jaw generating a higher LPP, it is equally uncertain what these effects were due to.

6.4 Overgeneralisation Effect

With the theory of fWHR being a cue of threat notwithstanding the results of the current study, a more probable explanation of the result in previous studies on fWHR, as well as the present study, is that they are due to the overgeneralisation effect. That is, that neutral faces with features resembling those of an angry face, e.g., low brows (Langner et al., 2010), could be overgeneralised and interpreted as the person having aggressive intentions and therefore be threatening (Said et al., 2009). Additionally, facial features that resemble those of angry faces might also cue strength and fighting ability and therefore be threatening (Sell et al., 2014). The overgeneralisation hypothesis has strong support from the finding that faces categorised as neutral (Engell, Haxby, & Todorov, 2007), were still rated in agreement to have more or less threatening personality traits (Said et al., 2009). Most importantly, the threatening faces' subtle facial structure correlated with the configuration of facial features seen in expressions of anger (Said et al., 2009). Hence, even subtle involvement of angry-looking features such as low brow height could through overgeneralisation effects and through being cues of strength explain the results in the present study as well as the results in most of the previous studies regarding the connections between fWHR and perceived threat. With this in mind, would the LPP be a good measure of perceived threat?

6.5 LPP and Threat Perception

Given that the facial ratings in the present study overlap nicely with the magnitude of the LPP in the 400-800ms window, the results support that LPP accurately indicates threat perception. However, as discussed above, the attenuation of the LPP in the late window (800-1500 ms) together with the uncertain separation of the LPP with the P300 component (Hajcak et al., 2010), makes it hard to relate the results to perceptual and underlying neurological mechanisms. This problem could perhaps be solved with future research devoted to developing research designs that could efficiently distinguish the LPP from the P300. Furthermore, investigating the relationship between threat perception and other early components such as the VPP, EPN and P300 might be fruitful given these components both unique and overlapping properties (Batty & Taylor, 2003; Duncan-Johnson & Donchin, 1977; Luck & Kappenman, 2011; Schupp, Öhman, et al., 2004). Doing so might create a more comprehensive way of indexing threat perception.

6.6 Limitations

The low sample size in the present study (facial ratings $N = 28$; ERP $N = 21$) reduces the validity of the results. However, large parts of the literature on correlations between high fWHR and threat perception (e.g., Carré et al., 2009; Carré, Morrissey, Mondloch, & McCormick, 2010; Efferson & Vogt, 2013; Geniole, Keyes, Mondloch, Carré, & McCormick, 2012; Hehman, Leitner, & Freeman, 2014; Short et al., 2012) found significant effects using sample sizes of 16 to 31 participants. Others, using an experimental design (e.g., Hehman et al., 2014; Zilioli et al., 2015) found significant effects with sample sizes between 36 to 40. With these numbers in mind, the present studies result are probably not only due to the sample size being too small to pick up on the effects of the fWHR.

Regarding the ERP measures, previous research has found significant results using sample sizes between 20 and 28 (Cuthbert et al., 2000; Schupp, Junghöfer, et al., 2004; Schupp, Öhman, et al., 2004; Schupp, Stockburger, et al., 2007). Therefore, the sample size of the present study is in the small end of the spectrum, but might still be adequate to generate enough power.

One fundamental issue with the design used in the present study was that participants had their electrodes gelled while they were doing the facial ratings. The gelling would most likely be a little distracting, perhaps making participants less sensitive to cues in the faces. However, if fWHR would be an evolutionarily evolved cue of threat, it would need to be robust enough to attain attention in far more stressful situations than electrode gelling. However, the cumulative effect of the researcher being in the same room, the gelling syringes poking the scalp while rating faces and the small sample size could have affected the power of the facial ratings. Although, as brow height had a significant effect size of $\eta_p^2 = .698$, the power of the facial ratings seems strong enough to pick up facial cues of threat.

7. **Conclusion**

The present study found support for that brow height has confounded the fWHR measure in most previous fWHR studies related to threat perception. Most importantly, due to overgeneralisation effects and possibly links to strength and fighting ability, it seems like brow height could explain most of the variance observed not only in this study but possibly in most other studies on fWHR related to threat perception. Whether brow height affects threat perception mostly due to overgeneralisation effects or due to it acting as a cue of strength is unclear. However, perception of someone expressing aggressive intent towards you, following from overgeneralisation effects, could arguably be seen as a more immediate and hence a greater threat than perceiving that a person is strong. Although, this might depend on the situation and ought to be investigated further. Moreover, if other features in neutral faces such as vertical eye narrowness and chin size also affect threat perception and strength judgement is still to be investigated. Furthermore, support for using LPP as an objective index of threat perception was found in the present study. However, the need for future research to resolve the ambiguity concerning the distinction between early LPP and P300 was clear.

One implication of fWHR being confounded by brow height and possibly generating most of the variance seen in previous fWHR studies could be that the findings of accurate inferences related to fWHR are instead related to brow height. Alternatively, other features, only correlating with fWHR and brow height, could have driven these inferences. Regardless, the validity of fWHR being a static facial feature evolutionary driven to cue threat can be questioned and deserves to be investigated with more rigorous control for other features.

The social consequences of judgments of facial features are huge. Judgments of competence predict elections (Todorov et al., 2005) and judgments of trustworthiness predict death sentences (Wilson & Rule, 2015). With appearance-based inferences today seemingly affecting who is put into power and who is executed, understanding the mechanisms underlying our appearance-based inferences is of great importance and every bit as intriguing as when Aristotle first commented on it. Regarding history, we should learn from past mistakes of leaning too much towards accounts of either accuracy or prejudice. Much like how we might mistake a twig for a snake, we might overgeneralise a neutral face with low brows for someone being angry and threatening. Also, much like how our perceptual bias towards snakes likely evolved for us to quickly and accurately detect snakes, our ability to make trait judgments of neutral faces could be linked to perceptual mechanism driven by evolution to detect cues of threat. The scientific prescription, then, is that both accounts should be equally explored and scrutinised, neither taken at face value.

Appendix A

The following procedure and settings were used when creating the stimuli in FaceGen. When you change one facial feature in FaceGen, other features will change with it. Therefore, the order of the settings matter. These instructions display settings and their order. Before any features were calibrated, all faces were set to ‘‘new”, ‘‘racial - any”, ‘‘male”, ‘‘skin texture 24” and ‘‘age” were set at approximately 25. Then, brows were calibrated, followed by cheekbones, then the jaw and lastly the eyes.

Variable Settings

Low brow: (1) set ‘‘brow ridge high/low” to 6, (2) set ‘‘brow ridge inner” to -6, and (3) set ‘‘brow ridge outer” to 4. High brow: use the average settings that follows from selecting ‘‘new”, ‘‘racial - any”, and ‘‘male”. Slim cheekbones: set “cheekbones - thin/wide” to -5. Wide cheekbones: set “cheekbones - thin/wide” to 4. Thin jaw: (1) set ‘‘jaw - thin/wide” to 2, and (2) set “jawline - concave/convex” to -2. Wide jaw: (1) set “jaw - thin/wide” to -5, and (2) set “jawline - concave/ convex” to 5. Opened eyes: use default settings. Closed eyes: set ‘‘eyelid - opened/closed” to 9.

Appendix B

Table 1.

Rejection rates after artifact detection.

Participant	Rejected trials per subject (%)
S1	14,5
S2	2,9
S3	0,4
S4	13,2
S5	4,6
S6	3,4
S7	4,8
S8	100,0
S9	5,1
S10	13,8
S11	8,3
S12	30,9
S13	8,2
S14	22,0
S15	21,9
S16	7,2
S17	8,6
S18	21,2
S19	21,9
S20	11,1
S21	4,9
S22	6,1
S23	8,3
S24	32,6
S25	4,0
S26	15,4
S27	0,3
S28	8,1
S29	36,8
S30	40,7

Note. Participants with grey marking on their rejected trials were removed altogether from the analysis.

References

Andersson, M. (1982). Female choice selects for extreme tail length in a widowbird. *Nature, 299*(5886), 818–820. doi:10.1038/299818a0

Archer, J. (2004). Sex differences in aggression in real-world settings: A meta-analytic review. *Review of General Psychology, 8*(4), 291–322. doi:10.1037/1089-2680.8.4.291.supp

Archer, J. (2009). Does sexual selection explain human sex differences in aggression? *Behavioral and Brain Sciences*, *32*(3–4), 249–266. doi:10.1017/S0140525X09990951

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

Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research*, *17*(3), 613–620. doi:10.1016/S0926-6410(03)00174-5

Berscheid, E. (1981). An overview of the psychological effects of physical attractiveness. In G. Lucker, K. Ribbens, & J. McNamara (Eds.), *Psychological aspects of facial form* (p. 1-23). Ann Arbor: University of Michigan Press.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, *14*(3), 519–521. doi:10.1017/S0140525X00071090

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. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1651), 2651–2656. doi:10.1098/rspb.2008.0873

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior: Research report. *Psychological Science*, *20*(10), 1194–1198. doi: 10.1111/j.1467-9280.2009.02423.x

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

Clutton-Brock, T. H., Guinness, F. E., & Albon, S. D. (1982). *Red deer: Behavior and ecology of two sexes*. University of Chicago press.

Clutton−Brock, T. H., & Harvey, P. H. (1977). Primate ecology and social organization. *Journal of Zoology, 183*(1), 1–39. doi:10.1111/j.1469-7998.1977.tb04171.x

Codispoti, M., Ferrari, V., & Bradley, M. M. (2006). Repetitive picture processing: autonomic and cortical correlates. *Brain Research*, *1068*(1), 213–220. doi:10.1016/j.brainres.2005.11.009

Codispoti, M., Ferrari, V., & Bradley, M. M. (2007). Repetition and event-related potentials: Distinguishing early and late processes in affective picture perception. *Journal of Cognitive Neuroscience*, *19*(4), 577–586. doi:10.1162/jocn.2007.19.4.577

Cohen, M. X. (2017). Where does EEG come from and what does it mean? *Trends in Neurosciences*, *40*(4), 208–218. doi:10.1016/j.tins.2017.02.004

Cunningham, M. R., Barbee, A. P., & Pike, C. L. (1990). What do women want? Facialmetric assessment of multiple motives in the perception of male facial physical attractiveness. *Journal of Personality and Social Psychology*, *59*(1), 61–72. doi:10.1037/0022-3514.59.1.61

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, *52*(2), 95–111. doi:10.1016/S0301-0511(99)00044-7

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London: Murray

Darwin, C. (1872). *The expression of the emotions in man and animals*. London: Murray

Deaner, R. O., Goetz, S. M. M., Shattuck, K., & Schnotala, T. (2012). Body weight, not facial width-to-height ratio, predicts aggression in pro hockey players. *Journal of Research in Personality*, *46*(2), 235–238. doi:10.1016/j.jrp.2012.01.005

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *13*, 9–21. doi:10.1016/j.jneumeth.2003.10.009

Dixson, B. J., & Vasey, P. L. (2012). Beards augment perceptions of men's age, social status, and aggressiveness, but not attractiveness. *Behavioral Ecology*, *23*(3), 481–490. doi:10.1093/beheco/arr214

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

Duncan−Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event−related potentials with subjective probability. *Psychophysiology*, *14*(5), 456–467. doi:10.1111/j.1469-8986.1977.tb01312.x

Efferson, C., & Vogt, S. (2013). Viewing men's faces does not lead to accurate predictions of trustworthiness. *Scientific Reports*, *3*(1), 1047. doi:10.1038/srep01047

Eldblom, H. (2018). *Facial width-to-height ratio as a cue of threat: An ERP study*. Retrieved from http://his.diva-portal.org/smash/record.jsf?pid=diva2%3A1218148&dswid=-7055

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: Saunders.

E-Prime [Computer software]. (2019). Retrieved from https://pstnet.com/products/e-prime/

FaceGen [Computer software]. (2018). Retrieved from https://facegen.com/index.htm

Fink, B., Grammer, K., Mitteroecker, P., Gunz, P., Schaefer, K., Bookstein, F. L., & Manning, J. T. (2005). Second to fourth digit ratio and face shape. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1576), 1995–2001. doi:10.1098/rspb.2005.3179

Fink, B., Neave, N., & Seydel, H. (2007). Male facial appearance signals physical strenght to women. *American Journal of Human Biology, 19*(1), 82–87. https://doi.org/10.1002/ajhb.20583

Fisher, R. A. (1930). *The genetic theory of natural selection*. Oxford: Clarendon Press

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

Foti, D., Hajcak, G., & Dien, J. (2009). Differentiating neural responses to emotional pictures: evidence from temporal-spatial PCA. *Psychophysiology*, *46*(3), 521–530. doi:10.1111/j.1469-8986.2009.00796.x

Geniole, S. N., Denson, T. F., Dixson, B. J., Carré, J. M., & McCormick, C. M. (2015). Evidence from meta-analyses of the facial width-to-height ratio as an evolved cue of threat. *PLoS ONE*, *10*(7), 132726. doi:10.1371/journal.pone.0132726

Geniole, S. N., Keyes, A. E., Mondloch, C. J., Carré, J. M., & McCormick, C. M. (2012). Facing aggression: Cues differ for female versus male faces. *PLoS ONE*, *7*(1), e30366. doi:10.1371/journal.pone.0030366

Geniole, S. N., & McCormick, C. M. (2015). Facing our ancestors: Judgements of aggression are consistent and related to the facial width-to-height ratio in men irrespective of beards. *Evolution and Human Behavior*, *36*(4), 279–285. doi:10.1016/j.evolhumbehav.2014.12.005

Gómez-Valdés, J., Hünemeier, T., Quinto-Sánchez, M., Paschetta, C., de Azevedo, S., González, M. F., … González-José, R. (2013). Lack of support for the association between facial shape and aggression: A reappraisal based on a worldwide population genetics perspective. *PLoS ONE*, *8*(1), e52317. doi:10.1371/journal.pone.0052317

Haines, S. E., & Gould, J. L. (1994). Female platys prefer long tails. Nature, *370*(6490), 512. doi: 10.1038/370512a0

Hajcak, G., Macnamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: An integrative review. *Developmental Neuropsychology*, *35*(2), 129–155. doi: 10.1080/87565640903526504

Hajcak, G., & Olvet, D. M. (2008). The persistence of attention to emotion: Brain potentials during and after picture presentation. *Emotion, 8*(2), 250–255. doi:10.1037/1528-3542.8.2.250

Hamilton, W., & Zuk, M. (1982). Heritable true fitness and bright birds: A role for parasites? *Science, 218*(4570), 384–387. doi:10.1126/science.7123238

Haselhuhn, M. P., Ormiston, M. E., & Wong, E. M. (2015). Men's facial width-to-height ratio predicts aggression: A meta-analysis. *PLoS ONE*, *10*(4), e0122637. doi:https://doi.org/10.1371/journal.pone.0122637

Hehman, E., Leitner, J. B., & Freeman, J. B. (2014). The face–time continuum: Lifespan changes in facial width-to-height ratio impact aging-associated perceptions. *Personality and Social Psychology Bulletin*, *40*(12), 1624–1636. doi:10.1177/0146167214552791

Isbell, L. A. (2006). Snakes as agents of evolutionary change in primate brains. *Journal of Human Evolution*, *51*(1), 1–35. doi:10.1016/j.jhevol.2005.12.012

Keil, A., Bradley, M., Hauk, O., Rockstroh, B., Elbert, T., & Lang, P. (2002). Large-scale neural correlates of affective picture processing. *Psychophysiology, 39*, 641–649. doi:10.1017/S0048577202394162

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior*, *20*(3), 165–182. doi:10.1007/bf02281954

Kramer, R. S. S. (2017). Sexual dimorphism of facial width-to-height ratio in human skulls and faces: A meta-analytical approach. *Evolution and Human Behavior*, *38*(3), 414–420. doi: 10.1016/j.evolhumbehav.2016.12.002

Kramer, R. S. S., Jones, A. L., & Ward, R. (2012). A lack of sexual dimorphism in width-to-height ratio in white european faces using 2D photographs, 3D scans, and anthropometry. *PLoS ONE, 7*(8), e42705. doi:10.1371/journal.pone.0042705

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. In P. J. Lang, R. F. Simons, & M. T. Balaban (Eds.), *Attention and orienting: Sensory and motivational processes* (p. 97-136). Hillsdale, NJ: Erlbaum.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the radboud faces database. *Cognition and Emotion*, *24*(8), 1377–1388. doi:10.1080/02699930903485076

Lefevre, C. E., Etchells, P. J., Howell, E. C., Clark, A. P., & Penton-Voak, I. S. (2014). Facial width-to-height ratio predicts self-reported dominance and aggression in males and females, but a measure of masculinity does not. *Biology Letters*, *10*(10), 20140729. doi:10.1098/rsbl.2014.0729

Little, A. C., Třebický, V., Havlíček, J., Roberts, S. C., & Kleisner, K. (2015). Human perception of fighting ability: Facial cues predict winners and losers in mixed martial arts fights. *Behavioral Ecology, 26*(6), 1470–1475. doi:10.1093/beheco/arv089

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 1-14. doi:10.3389/fnhum.2014.00213

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2. ed.). Cambridge, MA: MIT press.

Luck, S. J., & Kappenman, E. S. (Eds.). (2011). *The Oxford handbook of event-related potential components*. Oxford university press.

MacNamara, A., Foti, D., & Hajcak, G. (2009). Tell me about it: Neural activity elicited by emotional pictures and preceding descriptions. *Emotion, 9*(4), 531–543. doi:10.1037/a0016251

MacNamara, A., & Hajcak, G. (2010). Distinct electrocortical and behavioral evidence for increased attention to threat in generalized anxiety disorder. *Depression and Anxiety*, *27*(3), 234–243. doi:10.1002/da.20679

Manson, J. H., & Wrangham, R. W. (1991). Intergroup aggression in chimpanzees and humans. *Current Anthropology*, *32*(4), 369–390. doi:10.1086/203974

MATLAB [Computer software]. (2019) Retrieved from https://se.mathworks.com/help/matlab/release-notes-R2015a.html

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

Olofsson, J. K., & Polich, J. (2007). Affective visual event-related potentials: Arousal, repetition, and time-on-task. *Biological Psychology, 75*(1), 101–108. doi:10.1016/j.biopsycho.2006.12.006

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. doi:10.1073/pnas.0805664105

Ormiston, M. E., Wong, E. M., & Haselhuhn, M. P. (2017). Facial-width-to-height ratio predicts perceptions of integrity in males. *Personality and Individual Differences*, *105*, 40–42. doi:10.1016/J.PAID.2016.09.017

Parker, G. A. (1974). Assessment strategy and the evolution of fighting behaviour. *Journal of Theoretical Biology, 47*(1), 223–243. doi:10.1016/0022-5193(74)90111-8

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. doi:10.1037/0021-9010.88.5.879

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. doi:10.1080/00224499.2012.658924

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology*, *57*(1), 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? *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*, 93-95. doi:10.1098/rsbl.2003.0023

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. doi: 10.1016/j.evolhumbehav.2004.08.014

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

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Tiffany, I., & Lang, P. J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology*, *37*(2), 257–261. doi:10.1017/S0048577200001530

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Hillman, C. H., Hamm, A. O., & Lang, P. J. (2004). Brain processes in emotional perception: Motivated attention. *Cognition and Emotion*, *18*(5), 593–611. doi:10.1080/02699930341000239

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2004). The selective processing of briefly presented affective pictures: An ERP analysis. *Psychophysiology*, *41*(3), 441–449. doi:10.1111/j.1469-8986.2004.00174.x

Schupp, H. T., Stockburger, J., Codispoti, M., Junghofer, M., Weike, A. I., & Hamm, A. O. (2007). Selective visual attention to emotion. *Journal of Neuroscience*, *27*(5), 1082–1089. doi: 10.1523/JNEUROSCI.3223-06.2007

Schupp, H. T., Öhman, A., Junghöfer, M., Weike, A. I., Stockburger, J., Hamm, A. O., & Weike, I. (2004). The facilitated processing of threatening faces: An ERP analysis. *Emotion*, *4*(2), 189–200. doi:10.1037/1528-3542.4.2.189

Sell, A., 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., 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. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1656), 575–584. doi:10.1098/rspb.2008.1177

Sell, A., Hone, L. S. E., & Pound, N. (2012). The importance of physical strength to human males. *Human Nature, 23*(1), 30–44. doi:10.1007/s12110-012-9131-2

Shepherd, J. P., Gayford, J. J., Leslie, I. J., & Scully, C. (1988). Female victims of assault. *Journal of Cranio-Maxillofacial Surgery, 16*, 233–237. doi:10.1016/s1010-5182(88)80053-2

Short, L. A., Mondloch, C. J., McCormick, C. M., Carré, J. M., Ma, R., Fu, G., & Lee, K. (2012). Detection of propensity for aggression based on facial structure irrespective of face race. *Evolution and Human Behavior*, *33*(2), 121–129. doi:10.1016/j.evolhumbehav.2011.07.002

Soares, S. C., Lindström, B., Esteves, F., & Öhman, A. (2014). The hidden snake in the grass: Superior detection of snakes in challenging attentional conditions. *PLoS ONE*, *9*(12), e114724. doi:10.1371/journal.pone.0114724

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

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

Thornhill, R., & Gangestad, S. W. (1993). Human facial beauty. Human Nature, 4(3), 237–269. doi: 10.1007/bf02692201

Thornhill, R., & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences*, *3*(12), 452–460. doi: 10.1016/S1364-6613(99)01403-5

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Psychology: Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623–1626. 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. *Annu. Rev. Psychol*, *66*, 519–545. doi:10.1146/annurev-psych-113011-143831

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology*, *11*(4–5), 375–424. doi: 10.1016/0162-3095(90)90017-Z

Toscano, H., Schubert, T. W., & Sell, A. N. (2014). Judgments of dominance from the face track physical strength. *Evolutionary Psychology*, *12*(1), 1–18. doi:10.1177/147470491401200101

Třebický, V., Havlíček, J., Roberts, S. C., Little, A. C., & Kleisner, K. (2013). Perceived aggressiveness predicts fighting performance in mixed-martial-arts fighters. *Psychological Science*, *24*(9), 1664–1672. doi:10.1177/0956797613477117

Trivers, R. L. (1985) *Social Evolution.* Menlo Park, CA: Benjamin/Cummings

Twine, R. (2002). Physiognomy, phrenology and the temporality of the body. *Body & Society*, *8*(1), 67–88. doi:10.1177/1357034X02008001004

Valla, J. M., Ceci, S. J., & Williams, W. M. (2011). The accuracy of inferences about criminality based on facial appearance. *Journal of Social, Evolutionary, and Cultural Psychology*, *5*(1), 66–91.doi:10.1037/h0099274

Van Strien, J. W., Eijlers, R., Franken, I. H. A., & Huijding, J. (2014). Snake pictures draw more early attention than spider pictures in non-phobic women: Evidence from event-related brain potentials. *Biological Psychology*, *96*(1), 150–157. doi:10.1016/j.biopsycho.2013.12.014

Verdonck, A., Gaethofs, M., Carels, C., & DeZegher, F. (1999). Effect of low-dose testosterone treatment on craniofacial growth in boys with delayed puberty. *European Journal of Orthodontics*, *21*(2), 137–143. doi:10.1093/ejo/21.2.137

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. doi:10.1016/j.paid.2010.08.006

Weinberg, A., & Hajcak, G. (2010). Beyond good and evil: the time-course of neural activity elicited by specific picture content. *Emotion*, *10*(6), 767–782. doi:10.1037/a0020242

Weston, E. M., Friday, A. E., & Liò, P. (2007). Biometric evidence that sexual selection has shaped the hominin face. *PLoS ONE*, *2*(8), e710. doi:10.1371/journal.pone.0000710

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

Wilson, M., & Daly, M. (1985). Competitiveness, risk taking, and violence: The young male syndrome. *Ethology and Sociobiology*, *6*(1), 59–73. doi:10.1016/0162-3095(85)90041-X

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. doi:10.1002/ajhb.21219

Winkler, I., Haufe, S., & Tangermann, M. (2011). Automatic classification of artifactual ICA-components for artifact removal in EEG signals. *Behavioral and Brain Functions, 7*(1), 30. doi:10.1186/1744-9081-7- 30.

Zahavi, A. (1975). Mate selection: A selection for a handicap. *Journal of Theoretical Biology, 53*(1), 205–214. doi:10.1016/0022-5193(75)90111-3

Zebrowitz, L. A. (1997). *Reading faces: Window to the soul?* Boulder, CO: Westview Press.

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

Zilioli, S., Sell, A. N., Stirrat, M., Jagore, J., Vickerman, W., & Watson, N. V. (2015). Face of a fighter: bizygomatic width as a cue of formidability. *Aggressive Behavior*, *41*(4), 322–330. doi:10.1002/ab.21544

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Snakes in the grass. *Journal of Experiemntal Psychology: General*, *130*(3), 466–478. doi:10.1037/AXJ96-3445.130.3.466

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, *80*(3), 381–396. doi: 10.1037/0022-3514.80.3.381

Özener, B. (2012). Facial width-to-height ratio in a turkish population is not sexually dimorphic and is unrelated to aggressive behavior. *Evolution and Human Behavior*, *33*(3), 169–173. doi: 10.1016/j.evolhumbehav.2011.08.001