YAŞAR UNIVERSITY
GRADUATE SCHOOL

MASTER THESIS

HOW DOES AN EXTRAVERT OR A PSYCHOPATH LOOK LIKE? THE EFFECTS OF FACE FEATURES ON INDIVIDUALS' IMPRESSIONS OF PERSONALITY TRAITS

BÜŞRA ELİF YELBUZ

THESIS ADVISOR: ASSOC. PROF. SİNAN ALPER

PSYCHOLOGY PROGRAMME

PRESENTATION DATE: 08.06.2022

BORNOVA / IZMIR

JUNE 2022

ABSTRACT

HOW DOES AN EXTRAVERT OR A PSYCHOPATH LOOK LIKE? THE EFFECTS OF FACE FEATURES ON INDIVIDUALS' IMPRESSIONS OF PERSONALITY TRAITS

Yelbuz, Büşra Elif

MA, Psychology

Advisor: Assoc. Prof. Sinan Alper

June 2022

Research shows that individuals agree on certain facial features indicating certain fundamental traits such as valence and dominance. However, far less attention has been directed to how face perception influences trait judgements regarding more specific personality traits, namely, the Big Five and the Dark triad traits. Given that face-based trait impressions can have important social consequences, the current study aimed to contribute to this gap in the literature, by investigating how facial skin smoothness, facial baby facedness and facial masculinity affect individuals' impressions of strangers' Big Five and Dark Triad traits. An experimental study with 505 Turkish participants was conducted in which participants were asked to rate the personality traits of 16 faces that differed in their level of facial skin smoothness, baby facedness and masculinity. Namely, faces where manipulated such that there were two conditions (low versus high) for each of the three facial features (skin smoothness, baby facedness and masculinity) and it was assessed whether individuals' impressions of others' Big Five and Dark Triad traits differed in response to whether a face was low or high on each investigated facial feature (e.g., low skin smoothness versus high skin smoothness). Further, it was also explored whether the effect of facial skin smoothness and masculinity changed based on the face's gender. Results showed that faces high and low on facial skin smoothness, baby facedness and masculinity did indeed differ in their attributed Big Five and Dark Triad traits. Smoothed faces were assigned higher Dark Triad trait, Extraversion and lower Agreeableness ratings compared to blemished faces. Babyish faces were attributed higher Agreeableness and lower Dark Triad trait ratings compared to mature faces. Feminine faces were

attributed higher Agreeableness, Openness, Conscientiousness and lower Neuroticism and Dark Triad trait ratings compared to masculine faces. Finally, an interaction between gender and facial skin smoothness as well as gender and facial masculinity was found, such that the effect of skin smoothness and masculinity differed for male and female faces for most of the personality traits. Overall, these findings showed support for the broader assumption that individuals hold certain beliefs and stereotypes regarding how the Big Five and Dark Triad traits look like on a person's face. Possible explanations of the results, limitations and suggestions for future research were discussed.

Keywords: face perception, first impressions, impression formation, stereotypes, big five, dark triad, facial skin smoothness, facial masculinity, facial baby facedness

ÖZ

DIŞADÖNÜK VEYA PSİKOPATİ BİRİ NASIL GÖRÜNÜR? YÜZ ÖZELLİKLERİNİN İNSANLARIN KİŞİLİK İZLENİMLERİ ÜZERİNDEKİ ETKİSİ

Yelbuz, Büşra Elif
MA, Psychology
Advisor: Assoc. Prof. Sinan Alper
June 2022

Araştırmalar, belirli yüz özelliklerinin, güvenilirlik ve baskınlık gibi temel özelliklere işaret ettiğini göstermektedir. Ancak, yüz algısının daha spesifik kişilik özellikleri, yani Büyük Beş ve Karanlık üçlü özellikleri, ile ilgili yargılarını nasıl etkilediğine çok daha az odaklanılmıştır. Mevcut çalışma, yüze dayalı kişilik izlenimlerinin önemli sosyal sonuçları olabildiğini göz önünde bulundurarak, yüzün pürüzsüzlüğü, bebeksiliği ve maskülenliğinin; bireylerin, yabancıların Büyük Beş ve Karanlık Üçlü hakkındaki izlenimlerini nasıl etkilediğini araştırarak, literatürdeki bu boşluğa katkıda bulunmayı amaçlamıştır. Karanlık Üçlü. Beş yüz beş Türk katılımcıyla yapılan deneysel bir çalışmada, katılımcılardan pürüzsüzlük, bebeksilik ve maskülenlik seviyelerinde farklılık gösteren 16 yüzün kişilik özelliklerini derecelendirmeleri istendi. Yüzler, üç özellik için de (cilt pürüzsüzlüğü, bebeksilik ve maskülenlik) iki koşul olacak şekilde (düşük ve yüksek) manipüle edildi ve bireylerin, yüz özelliklerinin bu iki koşula bağlı olarak (örn: düşük pürüzsüzlük vs yüksek pürüksüzlük), Büyük Beş ve Karanlık Üçlü kişilik özelliklerine ilişkin izlenimlerinin, farklı olup olmadığı araştırıldı. Ek olarak, pürüzsüzlük ve maskülenlik etkisinin yüzün cinsiyetine göre değişip değişmediği de araştırıldı. Sonuçlar, yüz pürüzsüzlüğü, bebek yüzlülük ve maskülenliğin, yüksek ve düşük koşullar arasında Büyük Beş ve Karanlık Üçlü kişilik izlemilerinde gerçekten fark olduğunu gösterdi. Pürüzsüzleştirilmiş yüzlere, sivilceli yüzlere kıyasla, daha yüksek Karanlık Üçlü ve Dışadönüklük ve daha düşük Uyumluluk puanları verildi. Bebeksi yüzlere, olgun yüzlere kıyasla daha yüksek Uyumluluk ve daha düşük Karanlık Üçlü özellik derecelendirmeleri atfedildi. Feminen yüzlere, maskülen yüzlere kıyasla daha yüksek Uyumluluk, Açıklık, Sorumluluk ve

daha düşük Nevrotizm ve Karanlık Üçlü özellik derecelendirmeleri atfedildi. Son olarak, cinsiyet ile, cilt pürüzsüzlüğü ve yüz maskülenliği arasında bir etkileşim bulundu; cilt pürüzsüzlüğü ve maskülenlik etkisi, kişilik özelliklerinin çoğu için erkek ve kadın yüzlerinde farklılık gösterdi. Genel olarak, bu bulgular bireylerin, Büyük Beş ve Karanlık Üçlü özelliklerinin bir kişinin yüzünde nasıl göründüğüne ilişkin belirli inançlara ve stereotiplere sahip olduklarına dair daha geniş varsayımı desteklemektedir. Sonuçların olası açıklamaları, sınırlılıkları ve gelecek araştırmalar için öneriler tartışıldı.

Anahtar kelimeler: yüz algısı, ilk izlenimler, izlenim oluşumu, stereotipler, büyük beş, karanlık üçlü, yüz pürüzsüzlüğü, yüz maskülenliği, bebek yüzlülük

ACKNOWLEDGEMENTS

Firstly, I would like to thank my thesis advisor, Assoc. Prof. Sinan Alper, for not only all his invaluable feedback and advice throughout my studies, but also for being an inspiring academic with all his diligent and hard work. I am beyond grateful for having the opportunity to complete my thesis under his guidance and for being able to attend his lectures as he has both broadened my understanding of social psychology and open science and pushed me to always do better. Working with him was a true privilege. Additionally, a big thank you to my thesis jury members Prof. Adil Sarıbay and Assoc. Prof. Mert Teközel for helping me increase the quality of this thesis with all their constructive and valuable comments.

Secondly, I would like to thank my friends Elif Ecem Akten and Mustafa Kurses for tirelessly listening and contributing to all my thoughts and comments regarding my thesis and for always being there for me when I needed some extra motivation. They have supported me on this journey from day one, which I am eternally grateful for.

Finally, I would like to give, from the bottom of my heart, my endless thanks to my dearest family; my precious little brother Sina Senih Yelbuz, my compassionate mother Esma Yelbuz and my inspirational father Talat Mesud Yelbuz. I am beyond grateful to my brother for his constant support, entertainment, and technical knowledge, to my mother for her unconditional love, care, and attentiveness, and, last but not least, to my father for passing on his love for research to me, raising me to do everything I do with hard work and integrity and for his endless love and support.

Büşra Elif Yelbuz

June 2022

TEXT OF OATH

I declare and honestly confirm that my study, titled “HOW DOES AN EXTRAVERT OR A PSYCHOPATH LOOK LIKE? THE EFFECTS OF FACE FEATURES ON INDIVIDUALS’ IMPRESSIONS OF PERSONALITY TRAITS” and presented as a Master’s Thesis, has been written without applying to any assistance inconsistent with scientific ethics and traditions. I declare, to the best of my knowledge and belief, that all content and ideas drawn directly or indirectly from external sources are indicated in the text and listed in the list of references.

Büşra Elif Yelbuz

June 2022

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

It is commonly preached that you should not judge a book by its cover, but the reality is that we cannot stop ourselves from doing so. Individuals quickly and spontaneously form trait impressions, simply by looking at a person's facial appearance (Oosterhof & Todorov, 2008; Todorov et al., 2015; Willis & Todorov, 2006). Importantly, studies show that humans can agree on certain facial cues indicating certain traits such as valence (i.e., warmth, trustworthiness) and dominance (i.e., competence; Jones et al; 2021; Oosterhof & Todorov, 2008; Walker & Vetter, 2016). This points to the idea that there is a shared belief about what a trustworthy or dominant person looks like. However, far less is known about how face perception influences trait judgements regarding more specific personality traits, such as the Big Five and the Dark triad. This poses an interesting avenue of research: Similar to judgements of valence and dominance, do humans collectively associate certain face features with the big five and the dark triad personality traits? Moreover, what do people think an extravert, or a psychopath looks like? Accurate or not, judgements about strangers' personality can have great social consequences, ranging from who we choose to vote for (Todorov et al., 2005) to who we think it is acceptable to socially exclude (Rudert et al., 2017). Thus, it is crucial to work towards uncovering *how* individuals form judgments about others' personalities from just looking at their faces as well as what kind of stereotypes people rely on when making such judgements. The current proposal aimed to address this issue by experimentally manipulating a set of facial features and investigating their effects on peoples' beliefs about other's big five and dark triad traits. Drawing from the face perception literature, the effects of three previously studied facial features were assessed: Skin smoothness, baby facedness and masculinity. Specifically, faces were manipulated such that there were two conditions (low and high) for each of the three facial features and it was assessed whether individuals' impressions of others' big five and dark triad traits differed in response to whether a face was low or high on each facial feature (for example, low versus high on skin smoothness). With its novel findings, this pre-registered study provides

valuable information about individuals' beliefs and stereotypes regarding how people who score high and low on these specific personality traits look like.

CHAPTER 2 LITERATURE REVİEW

2.1. BACKGROUND OF FACE-BASED IMPRESSIONS

For quite some time now, cultural wisdom has warned us to not judge others by their superficial appearance. Not only does this suggest that judging people by their physical appearance will lead to mistaken first impressions, but it also assumes that we have a natural tendency to do so. As a matter of fact, this warning has a valid point. The literature on first impressions has collected ample evidence showing that people quickly and automatically form trait impressions, just from looking at a person's face (Jones et al., 2021; Oosterhof & Todorov, 2008; Todorov et al., 2015; Willis & Todorov, 2006). For example, Willis and Todorov (2006) found that very minimal exposure, as short as 100ms, to a stranger's face is enough for individuals to make specific trait inferences, such as how trustworthy, competent, or aggressive that person is. In fact, these inferences are made so rapidly that they do not appear to change much even after longer exposure to the faces (Willis & Todorov, 2006). This indicates that we have the tendency to make snap judgements about others' character just by looking at their face and that these judgements are not dependent on exposure time (Bar et al., 2006; Todorov et al., 2009; Todorov et al., 2010). Further, although the presence of some cultural differences should not be dismissed, there appears to be remarkable cross-cultural agreement, that even extends to populations of indigenous people, surrounding first impressions of fundamental traits such as how warm or dominant a person appears to be (Zebrowitz et al., 2012; Jones et al., 2021). Even infants and young children seem to respond to faces similar to how adults do (Cogsdill et al., 2014; Keating & Bai, 1986; Langlois et al., 1990; Montepare & Zebrowitz-McArthur, 1989).

Such universal findings point to the idea that our natural inclination of face-based impressions perhaps serve some adaptive function. In other words, it is likely that our tendency to judge a book by its cover supplied us with certain evolutionarily advantages in the past, even though cultural wisdom now warns against it. Indeed, the ecological approach to social perception, which is rooted in Gibson's (1979) theory of

object perception, proposes that people's facial appearance provides us with adaptive information on how to socially interact with them. For example, the 'cute' face of an infant evokes behavioural responses that are protective (Berry & McArthur, 1986; Zebrowitz, 1997) whereas an angry face elicits responses that are avoidant and defensive (Balaban, 1995; Marsh et al., 2005). Thus, we quickly and automatically generate certain assumptions on how to interact with others in the most appropriate way, based on their facial appearance.

On the first glance, this adaptive ability appears to be quite useful. However, even though this ecological approach assumes that our face-based impressions often turn out to be accurate, it also suggests that solely relying on certain facial information can result in biased perceptions due to overgeneralization effects (Zebrowitz, 1996; 1997; Zebrowitz & Montepare, 2006). Namely, facial cues that typically signal some specific personal trait (e.g., dominance) worth considering in our social interaction can influence our first impressions even when that person does not actually possess that personal characteristic, but just physically looks like someone who does (Zebrowitz & Collins, 1997). Within the ecological approach, such errors that result from overgeneralizations are assumed to be less maladaptive than those that could stem from failing to adjust our responses to people who vary in their facial appearance. In fact, this overgeneralization effect is just seen as a by-product of a broader cognitive mechanism, stimulus generalization, which is vital for humans' adaptive behaviour (Zebrowitz, 2017). Surely, without this ability we would not be able to have any expectations or presumptions about our social environment and the world would be a rather overwhelming place. Thus, overall, this natural inclination of face-based impressions appears to have been helpful in the past in appropriately guiding our social behaviour, to the point where it still continues to do so. Assessing what physical features are related to our face-based impressions is an important issue, because such impressions have significant consequences.

2.2. SOCIAL CONSEQUENCES OF FACE-BASED IMPRESSIONS

The literature is filled with practical examples of ways in which face-based impressions guide important social decisions and outcomes (Olivola et al., 2014b; Todorov et al., 2015). A large chunk of this work has focused on leadership selection and compensation (Todorov et al., 2015). Numerous studies have found that individuals tend to vote for faces that they judge to be more competent, and that such

faces are also more likely to win elections (Antonakis & Dalgas, 2013; Ballew & Todorov, 2007; Chen et al., 2014; Laustsen, 2014; Lenz & Lawson, 2011; Martin, 1978; Sussman et al., 2013; Todorov et al., 2005; Olivola & Todorov, 2010). It has also been shown that perceived dominance (Chen et al., 2014, Chiao et al., 2008, Little et al., 2007), sociability (Castelli et al., 2009) and threat (Mattes et al., 2010, Spezio et al., 2008) from faces predict electoral outcomes. Electoral outcomes are also predicted by how stereotypically Republican (Olivola et al., 2012) or politician-like (Olivola et al., 2014a) a candidate's face is perceived to be. From the business domain, studies show that competent and dominant looking CEOs receive higher salaries and are recruited by more successful companies (Graham et al., 2014; Rule & Ambady, 2008; 2009).

Further, studies have also pointed towards to importance of facial appearance in judgements of trust and guiltiness. For example, several lab studies have reliably demonstrated that, in strategic economic games, individuals are less willing to trust faces they perceive to be untrustworthy and more likely to cooperate with faces they see as trustworthy (Chang et al., 2010; Rezlescu et al., 2012; Schlicht et al., 2010; Stirrat & Perrett, 2010; Tingley, 2014; Van 't Wout & Sanfey, 2008). These findings hold even when the target's past behaviour signals that they can be trusted (Chang et al., 2010; Rezlescu et al., 2012). Worryingly, face-based impressions also can predict sentencing decisions. For example, an experimental study run by Porter and colleagues (2010) showed that untrustworthy-appearing defendants were more likely to be judged as guilty. For such faces, not only were participants more confident in their decisions but they also required less evidence to decide on whether the defendant was guilty or not (Porter et al., 2010). Yet most alarmingly, it appears that facial trustworthiness can even affect life or death decisions. Wilson and Rule (2015) found that untrustworthy-looking defendants are more likely to receive death sentences compared to trustworthy-looking defendants. Therefore, it is evident that facial appearance not only affects our judgements regarding who we see fit for certain social roles (e.g., a president, CEO, criminal etc.) it also influences our moral judgements. A final finding in support of this statement comes from Rudert et al (2017) who found that participants judged it as least morally acceptable to socially exclude a person that appeared warm-and-incompetent and most acceptable to socially exclude a cold-and-incompetent looking person.

All in all, these studies point to the practical importance of understanding *how* these quick impressions form as well as understanding which face features predict which trait inferences, since they ultimately end up translating themselves into our social behaviour.

2.3. THE VALENCE DOMINANCE MODEL

From a theoretical point of view, the most well-established account of how individuals evaluate faces in relation to social judgements has been Oosterhof and Todorov's Valence–Dominance Model (2008). Namely, Oosterhof and Todorov (2008) asked people to spontaneously rate faces on 13 different traits (aggressiveness, attractiveness, caringness, confidence, dominance, emotional stability, unhappiness, intelligence, meanness, responsibility, sociability, trustworthiness and weirdness). Importantly, Oosterhof and Todorov (2008) found that the ratings on these traits can be clustered into two factors. Consequently, they established a two-dimensional model of perception, valence and dominance, that is sufficient to predict individuals' social evaluations of faces. Valence can be understood as the degree to which a face is perceived as having harmful intentions and dominance can be described as the degree to which the face is perceived as having the capability to cause harm to the perceiver (Oosterhof & Todorov, 2008). In other words, the Valence–Dominance Model proposes that faces are primarily judged around two independent dimensions (i.e., valence and dominance) which form the basis of our first impressions (Oosterhof & Todorov, 2008; Todorov et al., 2015; Jones et al., 2021).

The Valence-Dominance Model has been successfully replicated in two western samples (Morrison et al., 2017; Wang et al., 2016), however in two Chinese samples (Sutherland et al., 2018; Wang et al., 2019) results were only half supportive. Namely, in both Chinese samples (Morrison et al., 2017; Wang et al., 2016) it was found that, while the participants social evaluations of faces were indeed guided by the valence dimension proposed by Oosterhof and Todorov (2008), there was no coherent finding of a corresponding dominance dimension. Rather, both studies (Morrison et al., 2017; Wang et al., 2016) described a second dimension, defined as capability, that was distinguished by rated intelligence. In light of these somewhat contradicting replication attempts, more recently, the Valence-Dominance Model has also been tested cross-culturally across 41 countries from 11 world regions (Jones et al., 2021). Importantly, Jones et al (2021) found that, when the original analysis strategy of

Oosterhof & Todorov (2008) was used to analyse data, results supported Valence-Dominance Model, therefore suggesting that these two dimensions are universally important for forming impressions. Notably though, the use of an alternative analysis revealed much less generalizability across world regions regarding to the Valence-Dominance Model (Jones et al., 2021). Nevertheless, when branching out from the first impressions literature, further support for this two-dimensional model comes from the stereotyping literature. Namely, the two dimensions put forward by Oosterhof and Todorov (valence and dominance; 2008) can be seen as conceptually corresponding to the two dimensions of the stereotype content model (Fiske et al., 2002; Fiske, 2018): Warmth and competence. Unsurprisingly, based on this foundational work, most of the research in the first impressions literature has focused on one or both of these dimensions (valence and dominance) when studying the social outcomes of face perception (Olivola et al., 2014b; Todorov et al., 2015).

2.4. FACE-BASED IMPRESSIONS OF WARMTH AND DOMINANCE

What physical features trigger such impressions that have important social consequences (section 2.2.)? There is now growing evidence showing that humans agree that certain facial cues indicate certain traits such as valence (warmth, trustworthiness) and dominance (competence; Jones et al; 2021; Oosterhof & Todorov, 2008; Todorov et al., 2008; Walker & Vetter, 2016; Todorov & Oh, 2021). For instance, Todorov and colleagues (2008) found that the facial features brow ridge (down/up), cheekbones (shallow/pronounced), chin (wide/thin) and nose sellion (shallow deep) were significant predictors of perceived trustworthiness. Namely, faces that had pronounced cheekbones, high inner eyebrows, wide chins, and shallow nose sellion were perceived to be more trustworthy compared to faces with shallow cheekbones, low inner eyebrows, thin chins and deep nose sellion. For dominance, facial masculinity has been found to be the most robust predictor of perceived dominance (Jones et al., 2010; Perrett et al., 1998; DeBruine et al., 2006; Boothroyd et al. 2007; Fink et al., 2007; Conway et al., 2009; Main et al. 2009). Zebrovitz (1997) pointed out that, while baby faced individuals have large eyes, high eyebrows, a smaller chin, round jaw, and high forehead, faces judged as dominant typically exhibit the opposite trend- small eyes, low brows, large chin, a more angular face, and a low forehead.

Researchers have even begun to compare *which* facial features (such as attractiveness, baby facedness, emotion resemblance, etc.) show the best predictive power over impressions of trustworthiness and dominance (Jaeger & Jones, 2021). Jaeger and Jones (2021) used machine learning techniques to test how well 28 important facial features are able to predict impressions of trustworthiness and dominance in a diverse set of 597 faces. All in all, they found that emotion resemblances were most predictive of both traits (Jaeger & Jones, 2021). Namely, resemblance to a happy expression was the strongest predictor for trustworthiness followed by perceived attractiveness and, for dominance, gender and resemblance to an angry expression were the strongest predictors (Jaeger & Jones, 2021).

The literature highlighted so far shows that extensive effort has been put into trying to understand how face perception influences inferences about two foundational traits, valence (i.e., trustworthiness, warmth) and dominance (i.e., competence), and how these inferences influence social behaviour. However, far less is known about how face perception influences trait judgements regarding more specific personality traits, such as the Big Five and the Dark triad. This poses an interesting avenue of research: Similar to judgements of valence and dominance, do humans collectively associate certain face features with the big five and the dark triad personality traits?

2.5. FACE-BASED IMPRESSIONS OF PERSONALITY TRAITS

The studies that exist so far, assessing individuals' face-based impressions of others' big five and dark triad traits, have largely focused on whether individuals can accurately predict others' personality traits (Alper et al., 2021; Holtzman, 2011; Borkenau et al., 2009; Penton-Voak et al., 2006; Little & Perett, 2007; Kramer & Ward, 2010; Walker & Vetter, 2016; Shiramizu et al., 2019). For example, Holtzman (2011) found that people were able to predict which faces scored higher on the Dark Triad traits (psychopathy, narcissism, and Machiavellianism) above chance level, though they performed better for female faces. Further, Alper and colleagues (2021) replicated and extended these findings in both an American (WEIRD) and Turkish (non-WEIRD) sample. Both American (WEIRD) and Turkish (non-WEIRD) participants were able to predict all Dark Triad personality traits above chance level as well as some of the Big Five traits, namely, extraversion (though only in females), agreeableness, and conscientiousness.

Within this line of research, the most popular choice of stimulus has been the use of computer graphic composites (i.e., averages; Alper et al., 2021; Holtzman, 2011; Penton-Voak et al., 2006; Little & Perett, 2007; Kramer & Ward, 2010; Shiramizu et al., 2019). These images are obtained through averaging the emotionally-neutral face images of people who have scored particularly high or particularly low on any interested trait, such that the final image comprises the average shape and colour of these faces. This method has been favoured due to the reasoning that composites (averages) allow the researcher to extract the defining characteristics of a group (exp: extraverts) while also eliminating the characteristics that make each face look individual (Rowland & Perrett, 1995; Penton-Voak et al., 2006). Overall, the findings of this body of work show that, by simply looking at such composite images of strangers, individuals are generally successful at predicting others' dark triad traits, though there are mixed results regarding accuracy concerning the Big Five traits (Alper et al., 2021; Holtzman 2011; Penton-Voak et al., 2006; Little & Perett, 2007; Kramer & Ward, 2010; Shiramizu et al., 2019).

2.6. ACURRACY VERSUS AGREEMENT

As a whole, the zero-acquaintance literature shows that individuals are at most moderately accurate in judging others' personalities from their faces (Walker & Vetter, 2016). In line with ecological approaches, this suggests that humans may indeed possess certain evolutionary adaptive mechanisms that allow us to infer someone's personality by simply looking at their face, but only to a certain degree (Alper et al., 2021). Nevertheless, as already mentioned, such judgements about strangers' personality can have great social consequences, regardless of whether they are accurate or not. Walker and Vetter (2016) argue that findings of high agreement between individuals (regarding trait impressions), as well as findings showing that individuals act upon such impressions, point towards a 'highly consequential socially shared reality'. Therefore, it is crucial to work towards better understanding the formation of personality judgments based on faces, regardless of whether these impressions prove to be accurate or not.

The predominant use of composite images within this literature has been helpful for assessing accuracy and to see whether an individual's idea of what, for example, an extraverted person looks like overlaps with how an extravert actually looks like. However, such methods leave the researcher with little experimental control

regarding what exactly participants base their judgements on. In other words, such an approach eliminates the opportunity to record and map out individuals' *ideas or stereotypes* around how an extravert looks. Thus, there is an important knowledge gap in the literature regarding whether humans agree on certain facial cues when forming impressions about strangers' big five and dark triad traits.

To my knowledge, only two studies (Walker & Vetter, 2016; Walker et al., 2018) so far have attempted to systematically manipulate facial cues and see whether there is agreement among participants regarding their Big Five personality judgements from faces. Walker and Vetter (2016) found that individuals do indeed exhibit strong agreement in their Big Five judgments from faces, concluding that individuals appear to have consensual beliefs on how, for example, an extraverted person looks like. However, their facial cue manipulation was based on the 2D space defined by the two basic dimensions of face evaluation (valence and dominance; Oosterhof & Todorov, 2008) and their results showed that the 2D space model was insufficient in explaining judgements of the Big Five dimensions. Thus, Walker and Vetter (2016) ultimately concluded that more research is needed to explore which facial features are used to make such personality judgments. More recently, the same group of researchers created the Basel Face Database in which 40 photographs of undergraduate students had been systematically manipulated to show reduced or enhanced values on the on the Big Five personality dimensions (Walker et al., 2018). Yet, their manipulations were quite subtle and solely based on shape and texture, therefore providing only limited information regarding which facial characteristics influence people's impressions or stereotypes of what, for example, an extravert looks like. Further, no study has assessed how directly manipulating certain facial features may affect impressions regarding others' dark triad traits. In light of these shortcomings, the current proposal aims to contribute to the first impressions and face perception literature by manipulating a set of facial features and directly investigating their effects on inferences about other's big five and dark triad traits.

2.7. MANIPULATED FACE FEATURES AND RESEARCH QUESTIONS

Drawing from previous studies that have investigated the effects of theory-driven facial features on more general trait judgements (such as trustworthiness and dominance; Jaeger & Jones, 2021), the effects of three previously studied facial

features were selected for assessment: Skin smoothness, baby facedness and masculinity. The reasoning behind the choice of these three facial features was based on (1) the fact that all three face features had previously been studied in the context of the overgeneralization hypothesis (Jaeger et al., 2017; Oosterhof & Todorov, 2008; Zebrowitz, 2017) and (2) the availability of previous findings showing their effects on trait judgements (such as trustworthiness and dominance) from which clear hypotheses could be derived regarding their effects on the big five and dark triad traits. Additionally, facial baby facedness and masculinity were specifically chosen because they are quite longstanding features of interest in the face perception literature (Oosterhof & Todorov, 2008; Perrett et al., 1998; Zebrowitz, 2017), whereas skin smoothness was specifically chosen because it is a more recently emerged feature of interest in the literature (Jaeger et al., 2017). The relevant literature for all three face features, from which clear and pre-registered hypotheses were derived from, are summarized in the next three sections.

2.7.1. Skin Smoothness

Research Question 1: How does facial skin smoothness affect face-based impressions of the Big Five and Dark Triad traits?

Formerly, Jaeger and colleagues (2017) investigated the effects of facial skin smoothness and blemishes on individuals' trait inferences (trustworthiness, competence, maturity). They found that, as predicted, blemished skin had a significant negative effect on trait ratings, such that faces with blemished skin were attributed lower trustworthiness, competence, maturity as well as lower attractiveness ratings compared to the baseline faces. On the other hand, skin smoothness only affected maturity and attractiveness ratings, such that faces with smoothed skin were seen as less mature and more attractive when compared to the baseline faces. Thus, Jaeger and colleagues (2017) concluded that the negative effect of blemished skin is more salient than the positive effect of smooth skin. In line with this, Tsankova and Kappas (2016) also found that the manipulation of skin smoothness affected explicit evaluations of trustworthiness, competence, attractiveness, and health. Thus, it could be expected that smooth faces are rated as higher on Agreeableness (i.e., trustworthiness) and Conscientiousness (i.e., competence), compared to blemished faces. Further, observer inferences of attractiveness have found to be correlated with inferences of Extraversion (Langlois et al., 2000) and Openness (Ćurković & Franc, 2010). Since attractiveness

is correlated with skin smoothness (Jaeger et at., 2017), smooth faces would be expected to be rated as higher on Extraversion and Openness, compared to blemished faces. Additionally, findings show that unattractive targets are assessed as more emotionally unstable (Ćurković & Franc, 2010), therefore blemished faces should be rated as higher on neuroticism, compared to smooth faces. Since the dark triad traits are associated with perceptions of untrustworthiness (Rogers et al., 2018; Gordon & Platek, 2009), skin smoothness should also be associated with the Dark Triad traits, such that blemished faces are rated as higher on Psychopathy and Machiavellianism, compared to smooth faces. However, for Narcissism, previous findings of a small but reliable positive narcissism–attractiveness correlation in observers' ratings (Holtzman & Strube, 2010; Rauthmann & Kolar, 2012) suggest that smooth faces should be rated higher on Narcissism, compared to blemished faces. Finally, previous research finding an interaction between facial attractiveness and gender of the face for impressions of Extraversion, Conscientiousness and Openness (Tartaglia & Rollero, 2015) suggests that the possible interaction between skin smoothness and gender on personality trait ratings is also worth exploring.

In sum it was expected that faces with blemished skin would be judged as having a higher score on more negative traits (Machiavellianism, Psychopathy and Neuroticism), whereas faces with smoothed skin would be judged as having a higher score on more positive traits (Extraversion, Agreeableness, Conscientiousness, Openness) and Narcissism. Further it was explored whether there was an interaction effect of gender and facial skin smoothness on impressions of Big Five and Dark Triad traits.

2.7.2. Baby Facedness

Research Question 2: How does facial baby facedness affect face-based impressions of the Big Five and Dark Triad traits?

Another popular face feature that has been proven to have noteworthy effects on impression formation is baby facedness (Zebrowitz & Montepare, 2008). Baby facedness, or sometimes referred to as babyish features, are characterized as faces with larger eyes, higher eyebrows, smaller nose bridges, rounder and less angular faces, thicker lips, a higher forehead, and a shorter chin (Montepare & Zebrowitz, 1998; Zebrowitz, 1997; Zebrowitz et al., 2003). Previous work shows that individuals with such babyish features are perceived to have childlike traits, namely they are seen as

naïve, submissive, weak, warm, and honest (Montepare & Zebrowitz, 1998; Zebrowitz & Montepare, 1992). Thus, such babyish features could also be expected to influence individuals' impressions of strangers' personality traits. Baby facedness could prove to be especially relevant for impressions regarding the dark triad traits, since the dark triad traits have previously been found to be associated with negative impressions (Rauthmann & Kolar, 2012) and perceptions of lower trustworthiness (Rogers et al., 2018; Gordon & Platek, 2009), although the association with trustworthiness was found most strongly for psychopathy (Gordon & Platek, 2009). Further, individuals with babyish features have been found to be avoided for assigning mentally challenging tasks and leadership positions but favoured for jobs that require warmth and friendliness (Montepare & Zebrowitz, 1998). Thus, it can be concluded that baby facedness is overall associated with being warm but incompetent. Drawing on these mentioned studies, one would expect that mature faces would be rated as higher on Conscientiousness, Psychopathy, Machiavellianism, Narcissism, compared to babyish faces. Further, babyish faces should be rated as higher on Agreeableness, compared to mature faces. On the other hand, how babyish features affect perceptions of Openness, Extraversion and Neuroticism is a rather explorative question.

In sum it was expected that babyish faces would be judged as having a higher score on agreeableness, whereas more mature faces would be judged as having a higher score on Conscientiousness, Psychopathy, Machiavellianism and Narcissism. Further, it was explored how babyish features affected impressions of Openness, Extraversion and Neuroticism.

2.7.3. Masculinity

Research Question 3: How does facial masculinity affect face-based impressions of the Big Five and Dark Triad traits?

Finally, the last facial feature that will be assessed is facial masculinity/femininity, which has previously been found to be related to impressions of dominance (Boothroyd et al. 2007; Buckingham et al., 2006; Oosterhof & Todorov 2008). Further, Perrett et al (1998) reported that masculine faces are generally perceived to be less trustworthy when compared with feminine faces. In line with these results, Walker and Wänker (2017) also showed that masculine-looking faces were seen as colder and more competent than feminine-looking faces. Further, Walker and Wänker (2017) found that this effect of facial masculinity/femininity was actually

more pronounced for atypical (i.e., masculine-looking women and feminine-looking men) faces than for typical (i.e., masculine-looking men and feminine-looking women) faces. Based on these findings it seems plausible to assume that facial masculinity and femininity would also provide certain cues regarding others' personality traits. Namely, since the dark triad traits are associated with perceptions of untrustworthiness (Rogers et al., 2018; Gordon & Platek, 2009), one would expect that masculine-looking faces (both for men and women) would be rated as higher on Psychopathy, Machiavellianism and Narcissism, compared to feminine faces. Further, drawing on Walker and Vetter's (2016) findings showing that perceptions of agreeableness and openness to experience strongly overlap with trustworthiness and have a strong negative correlation with dominance, feminine faces should be judged as having a higher score on agreeableness and openness, compared to masculine faces. Since neuroticism, extraversion, and conscientiousness have been found go beyond the 2D space of trustworthiness and dominance (Walker & Vetter, 2016), the influence of facial femininity and masculinity on these traits is an exploratory question worth looking into. Further, inspired by the previously mentioned findings of Walker and Wänker (2017), the effect of facial masculinity/femininity being more pronounced for atypical (i.e., masculine-looking women and feminine-looking men) faces, it can be explored whether there is an interaction effect between the gender and femininity/masculinity of faces.

Thus, in sum, it was expected that masculine faces would be judged as having a higher score on the Dark Triad traits (Psychopathy, Machiavellianism and Narcissism), whereas feminine faces would be judged as having a higher score on agreeableness and openness. Additionally, it was explored how facial masculinity affected impressions of neuroticism, extraversion, and conscientiousness and whether there was an interaction effect of gender and facial femininity/masculinity on impressions of Big Five and Dark Triad traits.

2.8. OVERALL AIM AND HYPOTHESES

All in all, the current pre-registered study looked at the effects of three facial features (skin smoothness, baby facedness and masculinity/femininity), known to be influential on impression formation, in relation to individuals' ideas and stereotypes concerning the Big Five and Dark Triad traits. When studying first impressions, previous work has mainly focused on the effect of facial features on inferences of

valence and dominance (Oosterhoff & Todorov, 2008). Impressions regarding more specific traits (e.g., Big Five and Dark Triad traits) have received less attention in the face perception literature. The studies that do exist have (1) mostly been concerned with accuracy (typically via composite images), leaving little experimental control, or (2) employed only subtle manipulations, producing only very limited insight to how face features affect impressions of the Big Five and Dark Triad traits. To address this gap in the literature, faces were manipulated such that there were two conditions (low and high) for each of the three facial features and it was assessed whether individuals' impressions of others' big five and dark triad traits differed depending on whether a face is low or high on each facial feature (for example: Low and high on skin smoothness).

Based on the summarized literature, the hypotheses are as follows:

1. Skin smoothness Hypotheses (H1):

Faces with blemished skin (i.e., low level skin smoothness) will be judged as having a higher score on more negative traits (Machiavellianism, Psychopathy and Neuroticism), compared to faces with smoothed skin (i.e., high level skin smoothness) and faces with smoothed skin (i.e., high level skin smoothness) will be judged as having a higher score on more positive traits (Extraversion, Agreeableness, Conscientiousness, Openness) and Narcissism, compared to faces with blemished skin (i.e., low level skin smoothness).

2. Baby Facedness Hypotheses (H2):

Faces with babyish features (i.e., high level baby facedness) will be judged as having a higher score on agreeableness, compared to faces with mature features (i.e., low level baby facedness) and faces with mature features (i.e., low level baby facedness) will be judged as having a higher score on Conscientiousness, Psychopathy, Machiavellianism and Narcissism, compared to faces with babyish features (i.e., high level baby facedness).

3. Masculinity/Femininity Hypotheses (H3):

Faces with masculine features (i.e., high level masculinity) will be judged as having a higher score on the Dark Triad traits (Psychopathy, Machiavellianism and Narcissism), compared to faces with feminine features (i.e., low level masculinity) and faces with feminine features (i.e., low level masculinity) will be judged as having a higher score on agreeableness and openness, compared to faces with masculine features (i.e., high level masculinity).

Additionally, it will be explored (1) whether the influence of facial skin smoothness on impressions of personality traits differs for male and female faces, (2) how facial baby facedness influences impressions of Openness, Extraversion and Neuroticism, (3) how facial masculinity influences impressions of Extraversion, Neuroticism and Conscientiousness and, finally, (4) whether the influence of facial masculinity on impressions of personality traits differs for male and female faces. Thus, the explorative questions of this study are as follows:

Exploratory Question 1: Is there an interaction effect of gender (male versus female) and facial skin smoothness (high versus low) on impressions of personality traits?

Exploratory Question 2: How does facial baby facedness (high versus low) affect impressions of Openness, Extraversion and Neuroticism?

Exploratory Question 3: How does facial masculinity (high versus low) affect impressions of neuroticism, extraversion, and conscientiousness?

Exploratory Question 4: Is there an interaction effect of gender (male versus female) and facial masculinity (high versus low) on impressions of personality traits?

CHAPTER 3 PILOT STUDY

Prior to the main study, a pilot study was conducted to assess whether the faces manipulated on skin smoothness, baby facedness and masculinity/femininity did indeed differ in perceived skin smoothness, baby facedness and masculinity. Namely, to assess whether the face manipulations were successful, a within-subjects design with the independent variable manipulation level (high vs. low) and the dependent variables skin smoothness, masculinity, and baby facedness was utilized.

Power analysis revealed a minimum of 27 participants to ensure an 80% chance of detecting a medium effect (d = .05) at an alpha level of .05. A total of 40 participants were recruited, however only 21 participants completed the study and were included in the analyses. The study was conducted through Qualtrics, and the participants consisted of a convenience sample.

3.1. MATERIALS

Initially, a fourth facial feature was planned to be assessed: Facial attractiveness. However, upon conducting the pilot study, facial attractiveness was eliminated as a fourth face feature category (further explained in the discussion section 3.5.). Nevertheless, the initial face stimuli were created with a fourth facial feature (facial attractiveness) included. A total of 128 (4 facial features x 2 manipulation levels x 8 traits x 2 genders = 128) manipulated images were generated for this study. The original images were taken from the Bogazici Face Database (Saribay et al., 2018). 32 male (4 facial features x 8 traits = 32) and 32 female individuals with a frontal gaze and a neutral facial expression were selected, resulting in 64 distinct face identities. Prior to the manipulations, drawing on Saribay and colleagues (2018) validation study of the Bogazici face database, it was ensured that there were no gender differences in perceived attractiveness, $t(62) = 0.71$, $p = .48$, 95% C.I. [-0.86, 0.18], $d = .18$, perceived dominance, $t(62) = 0.44$, $p = .66$, 95% C.I. [-0.24, 0.18], $d = .11$, perceived trustworthiness, $t(62) = 1.50$, $p = .14$, 95% C.I. [-0.51, 0.36], $d = .38$, and perceived

Turkishness, $t(62) = 0.44$, $p = .66$, 95% C.I. [-0.19, 0.29], $d = .11$, of the individuals. However, there were significant gender differences in masculinity, $t(62) = -6.65$, $p < .001$, 95% C.I. [-1.32, -0.71], $d = 1.66$; $F_{welch}(1, 61.98) = 44.22$, $p < .001$, and femininity, $t(62) = 13.75$, $p < .001$, 95% C.I. [1.29, 1.73], $d = 3.43$; $F_{welch}(1, 49.07) = 189.07$, $p < .001$. It was also ensured that, prior to the manipulations, there were no significant differences in attractiveness, $F(3,60) = .797$, $p = .50$), dominance, $F(3, 63) = .786$, $p = .58$, masculinity, $F(3, 63) = .597$, $p = 62$, femininity, $F(3, 63) = .109$, $p = .95$, trustworthiness, $F(3, 63) = 2.02$, $p = .12$, and Turkishness, $F(3, 63) = .541$, $p = .66$, between the groups of faces chosen for each face feature category (Skin smoothness, Attractiveness, Baby facedness and Masculinity).

3.2. FACE FEATURE MANIPULATIONS

From the final set of 128 generated faces, 16 face identities were manipulated in relation to skin smoothness resulting in 16 high and 16 low skin smoothness faces (for a total of 32 faces), 16 face identities were manipulated in relation to baby facedness resulting in 16 high and 16 low baby facedness faces (for a total of 32 faces), 16 face identities were manipulated in relation to masculinity resulting in 16 high and 16 low masculinity faces (for a total of 32 faces) and, finally, 16 face identities were manipulated in relation to attractiveness resulting in 16 high and 16 low attractiveness faces (for a total of 32 faces). As already mentioned, since facial attractiveness was later eliminated from the main study (further explained in the discussion section 3.5.), the details of the face manipulations and pilot study results regarding facial attractiveness are not reported.

3.2.1. Skin Smoothness Faces

Sixteen face identities were manipulated for skin smoothness. Two different versions of each face were created (for a total of 32 stimuli): a smoothed (i.e., high level) version and a blemished (i.e., low level) version. The images were manipulated using the GNU Image Manipulation Program (GIMP) version 2.10.30.

The skin smoothness manipulation was replicated from previous work (Jaeger et al. 2017), such that for the smoothed version, firstly a duplicate layer of the face image was created, and a 4.0 Gaussian blur filter was applied to the bottom layer. Next, the target areas (cheeks, chin, forehead) of the top layer were erased so that the smoothed bottom layer appeared for those areas. As Jaeger et al (2017) mentioned,

this ensured that only the target of the face would appear smoothed, while everything else stayed the same.

For the blemished version of the face, firstly, images of individuals with blemishes were downloaded from Google Images. Next, these blemishes were cropped and placed to different areas of the target faces while making sure that the faces still looked realistic. The faces had a black background, and the images were 300 x 300 pixels. Due to the copyrights associated with the original images (Saribay et al., 2018), no exemplar photos of the manipulated faces could be presented within the scope of this thesis. However, all materials and data can be provided upon request for any research purposes.

3.2.2. Baby Facedness Faces

Sixteen face identities were manipulated for baby facedness. Two different versions of each face were created (for a total of 32 stimuli): a babyish (i.e., high level) version and a mature (i.e., low level) version. The images were manipulated using FaceGen.

To create babyish faces, firstly higher brows were created by positioning the brow ridge (high/low) dimension slider to -3.00 SDs. Next, a narrower chin was achieved by positioning the chin (wide/thin) dimension slider to +3.00 SDs, larger eyes were created by positioning the eyes (small/large) dimension slider to +2.00 SDs and a rounder face was achieved by positioning the face (round/gaunt) dimension slider to -3.00 SDs. Finally, a smaller nose was achieved by positioning the nose bridge (small/large) dimension slider to -2.00 SDs and the nostril tilt (down/up) dimension to 0.00 SDs. For all 16 face identities, the high baby facedness manipulation was applied in this exact order. While the final set of values (in SDs) for these mentioned target areas were somewhat different for each face (due to FaceGen adjusting the SD values to accommodate the other characteristics of the face), the direction (negative/positive) of the target values was the same for all faces. To create mature faces, simply the age dimension slider was positioned to 40.00 SDs for all 16 identities, which, simply put, alters the face features associated with age to appear older.

3.2.3. Masculinity Faces

Sixteen face identities were manipulated for masculinity. Two different versions of each face were created (for a total of 32 stimuli): a masculine (i.e., high

level) version and a feminine (i.e., low level) version. The images were manipulated using FaceGen Modeller.

For the masculine version of the faces, firstly the gender dimension slider was positioned to -1.00 SDs for both male and female faces. In FaceGen, the negative values on the gender dimension correspond to masculinized versions of the initial face, which, simply put, alters the face features associated with gender to appear more masculine. Next, due to this manipulation darkening the skin tone of female faces, the skin tone of female faces was re-adjusted via the skin shade and skin flush sliders. This adjustment was made based on the reference values provided by FaceGen for the femininized version of the female faces.

For the feminine version of the faces, firstly the gender dimension slider was positioned to +1.00 SDs for both male and female faces. In FaceGen, the positive values on the gender dimension correspond to femininized versions of the initial face, which, again simply put, alters the face features associated with gender to appear more feminine. Next, due to this manipulation lightening the skin tone of male faces, the skin tone of male faces was re-adjusted via the skin shade and skin flush sliders. This adjustment was made based on the excel values provided by FaceGen for the masculinized version of the male faces. These skin tone adjustments were applied so that the difference the two manipulation levels (masculine vs. feminine) were solely based on differences in facial masculinity and femininity rather than any underlying biases surrounding skin tone.

3.3. PROCEDURE

Prior to conducting the pilot study all hypotheses, design and planned analyses were pre-registered at https://osf.io/8c7jw.

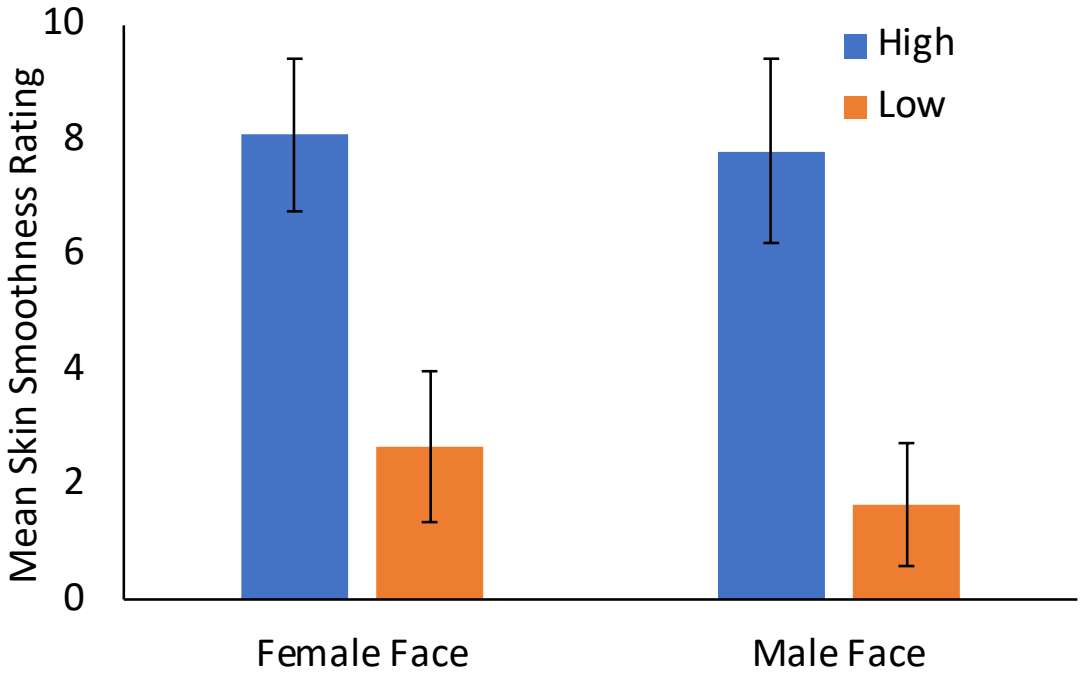
After providing written informed consent, participants rated the smoothness, attractiveness, baby facedness and masculinity of the 128 faces on a scale from 1 (not *smooth, attractive, baby faced* or *masculine* at all) to 9 (extremely *smooth, attractive, baby faced* or *masculine*). Specifically, participants were shown two faces, the low level and high level version of the face, at a time and asked to report how smooth, attractive, baby faced or masculine/feminine they thought the person in the photo was, depending on the which face feature was manipulated in that set of faces (for example, if masculinity was manipulated they were only asked to report the face's masculinity). The order of presentation for sets of faces was randomized for each participant using

the block randomizer function on Qualtrics. Further, for each set of faces, half of the participants saw the high manipulation level on the right side of the screen while the other half saw it on the left side of the screen, which was ensured via the evenly present elements function on Qualtrics. Unfinished surveys were not recorded, and it took an average of 20 mins to finish the study.

3.4. RESULTS

3.4.1. Skin Smoothness Results

Paired samples t-tests were conducted for each set of faces manipulated on skin smoothness. For all manipulated faces, results showed a significant difference between the smoothed skin (i.e., high level) face and blemished skin (i.e., low level) face in skin smoothness ratings ($p < .001$; see Appendix A). Thus, showing that the skin smoothness manipulation was successful. Figure 3.1. shows the mean perceived skin smoothness ratings for the high and low manipulation level faces used in the main study.

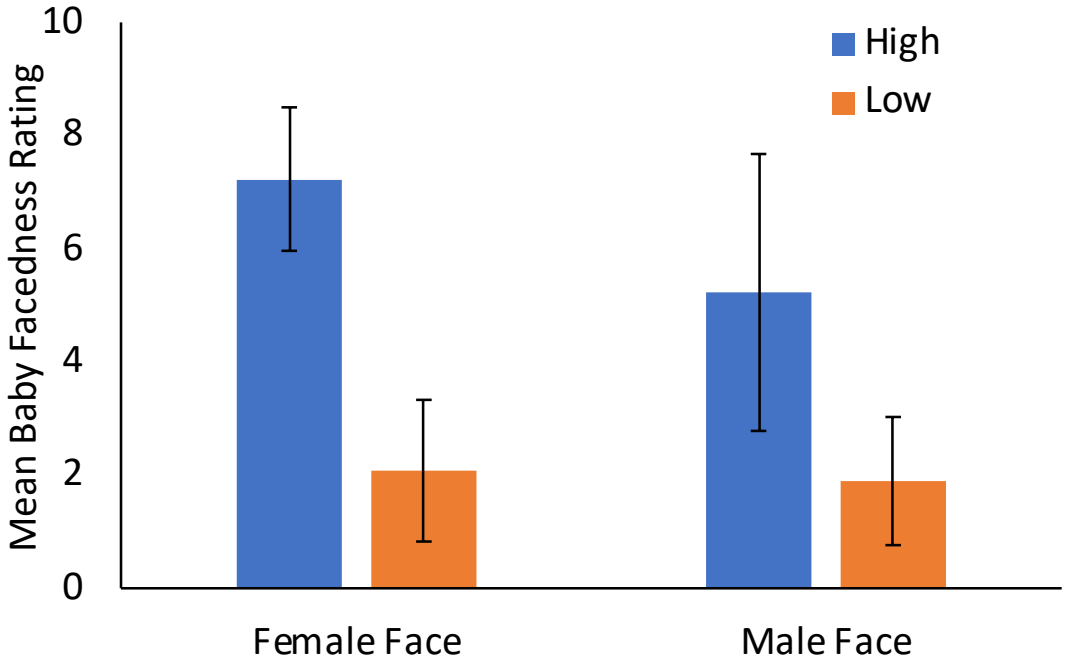


Note. Results of the face pairs used in the main study. Error bars represent SDs.

Figure. 3.1. Perceived Skin Smoothness for Faces High and Low in Skin Smoothness

3.4.2. Baby facedness Results

Paired samples t-tests were conducted for each set of faces manipulated on baby facedness. For all manipulated faces, results showed a significant difference between the babyish (i.e., high level) face and mature (i.e., low level) face in baby facedness ratings ($p < .001$, except for one set which was $p = .002$; see Appendix A). Thus, showing that the baby facedness manipulation was successful. Figure 3.2. shows the mean perceived baby facedness ratings for the high and low manipulation level faces used in the main study.

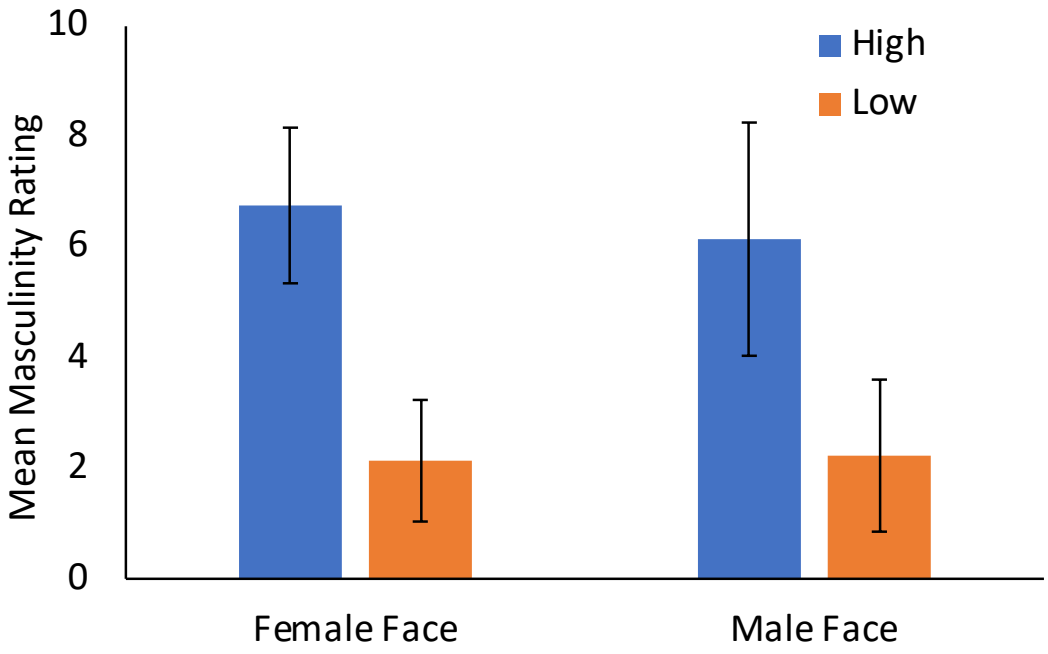


Note. Results of the face pairs used in the main study. Error bars represent SDs.

Figure. 3.2. Perceived Baby Facedness for Faces High and Low in Baby Facedness

3.4.3. Masculinity Results

Paired samples t-tests were conducted for each set of faces manipulated on masculinity. For all manipulated faces, results showed a significant difference between the masculine (i.e., high level) face and feminine (i.e., low level) face in masculinity ratings ($p < .001$; see Appendix A). Thus, showing that the masculinity manipulation was successful. Figure 3.3. shows the mean perceived masculinity ratings for the high and low manipulation level faces used in the main study.



Note. Results of the face pairs used in the main study. Error bars represent SDs.

Figure. 3.3. Perceived Masculinity for Faces High and Low in Masculinity

3.5. DISCUSSION

Based on the pilot study's results, the high and low levels of faces manipulated on skin smoothness, baby facedness and masculinity did indeed differ in perceived levels of skin smoothness, baby facedness and masculinity/femininity. Namely, smoothed skin faces were attributed higher ratings of skin smoothness compared to blemished skin faces, babyish faces were attributed higher ratings of baby facedness compared to mature faces and masculine faces were attributed higher ratings of masculinity compared to feminine faces.

Initially, all 128 faces were planned to be utilized in the main study so that (1) four facial features were assessed in total and (2) there would be a different face identity for each personality trait (psychopathy, narcissism, Machiavellianism, agreeableness, openness, conscientiousness, extraversion). However, from the pilot study it was observed that rating 128 faces made the study too long in terms of duration. Thus, to preserve participation and attention, two main alterations were made: (1) Attractiveness was eliminated as a facial feature due to its previously found positive correlation with skin smoothness (Fink et. al., 2001; Tsankova & Kappas, 2016; Jaeger et al., 2017) suggesting that keeping both facial features (attractiveness

and skin smoothness) could add unnecessary length to the main study. Further, (2) it was decided that participants would rate all eight personality traits on the same face to shorten the duration of the study. Namely, from these obtained pairs of manipulated faces, one male and one female pair of faces was chosen for each of the three face feature categories, resulting in a total of 12 faces (3 face features x 2 manipulation levels x 2 genders) to be rated by participants on all eight personality traits in the main study. The pairs of faces used in the main study were namely the face pairs that showed one of the highest mean differences between their ratings on the relevant face feature (skin smoothness, baby facedness and masculinity). As a final adjustment for the main study, the background of the faces was changed from black to white. All relevant adjustments were updated on the pre-registration (https://osf.io/8c7jw) prior to data collection for the main study.

CHAPTER 4 METHOD

4.1. PARTICIPANTS

Power analysis, conducted through G*Power, revealed a minimum of 163 participants to ensure an 80% chance of detecting a small effect (f = .10) for the interaction term (Facial Feature X Manipulation level) at an alpha level of .05. Thus, it was aimed to recruit at least 163 Turkish participants for this study. To be eligible participants had to be at least 18 years old and speak Turkish as their first language. Data collection occurred through various social media platforms (Twitter, Facebook, and Instagram). Participants were excluded from the final data analysis if they failed to finish the study or did not provide their informed consent. A total of 505 participants fully completed the study and were included in the analyses. The participants age ranged from 18 to 72 (M=30.8, SD = 9.63) and 76.44% (N = 386) of the participants were female. Further details can be found in the demographics table (Table 4.1.).

Table 4.1. Demographics Table

Demographic Variables	Categories	N	Percent	Mean	SD
Gender					
	Male	118	23.37		
	Female	386	76.44		
	Other	1	0.20		
Age (years old)		505		30.80	9.63
	18-24	138	27,33		
	25-34	230	45,54		
	35-44	82	16,24		
	45-54	41	8,12		
	55-64	10	1,98		
	65+	4	0,79		
Education					
	Elementary School	2	0.40		
	Middle School	3	0.59		
	High School	21	4.16		
	2-year Bachelor's Degree	18	3.56		
	4-year Bachelor's Degree	229	43.35		
	Master's Degree	160	31.68		
	Doctorate	72	14.26		

4.2. DESIGN

A 3 (Facial feature: skin smoothness, baby facedness and masculinity/femininity) x 2 (Manipulation levels: Low and high) x 2 (Gender: Male and Female) within-subjects design was used. Participants were shown a total of 12 faces (differing in manipulated facial features, manipulation levels and gender) and asked to give personality trait ratings for each face. Participant's ratings of how high they thought each face would score on the personality traits (Psychopathy, Machiavellianism, Narcissism, Extraversion, Neuroticism, Agreeableness, Conscientiousness and Openness) comprised the dependent variables.

4.3. MATERIALS

A total of 12 (3 facial features x 2 manipulation levels x 2 genders = 12) faces were used as the materials for the main study. As already mentioned, the original images were taken from the Bogazici Face Database (Saribay et al., 2018). 3 male and 3 female individuals with a frontal gaze and a neutral facial expression were selected, resulting in a total of 6 distinct face identities. Next these face identities were manipulated on FaceGen Modeller and GIMP in relation to skin smoothness, baby facedness and masculinity, such that there were 2 versions (Manipulation levels: Low and High) of each face identity, resulting in a final total of 12 images (explained in more detail in chapter 3 and 4). The faces had a white background, and the images were 300 x 300 pixels. Again, due to the copyrights associated with the original images (Saribay et al., 2018), no exemplar photos of the manipulated faces could be presented within the scope of this thesis. However, all materials and data can be provided upon request for any research purposes.

4.4. DEPENDENT VARIABLES

There was a total of eight dependent variables: Psychopathy, Machiavellianism, Narcissism, Extraversion, Neuroticism, Agreeableness, Conscientiousness and Openness ratings. Namely, for each of the 12 faces (varying in manipulated face feature, manipulation level and gender), participants gave a rating for each personality trait on a 9-point Likert scale ranging from 1 (not [trait] at all) to 9 (extremely [trait]). The mean ratings of each personality trait for each of the 12 faces

were computed and compared with respect to the independent variables facial feature, manipulation level and gender.

4.5. PROCEDURE

Firstly, participants were asked to provide their informed consent and notified that participation was voluntary and that they could choose to leave the study at any point in time. Next, participants were informed that they would be shown a total of 12 faces, and that they should try to guess the personality traits of each person by simply looking at their face. Thereafter, participants were presented with one face at a time and asked to give eight personality trait ratings for each face. Specifically, a simple definition of each personality trait was provided under every personality trait rating question and participants were asked 'How [trait] do you think this person is? 1 = not [trait] at all, 9 = extremely [trait]'. The order of trials (i.e., faces) were randomized for each participant using the Qualtrics' randomizer function. Further, the order of personality trait rating questions was randomized for each face using Qualtrics' randomizer function. After completing their personality trait ratings for all 12 faces, participants were asked to provide their demographic details (age, gender, education level, socioeconomic status, religiosity and political orientation). Finally, participants were thanked for their participation.

4.6. DATA ANALYSES

To test the hypotheses of the study, for each personality trait, a three-way within-subjects ANOVA, with one within-subjects factor of *Facial Feature* (three levels: skin smoothness, baby facedness and masculinity/femininity), another within-subjects factor of *Manipulation Level* (two levels: high and low) and a final within-subjects factor of *gender* (two levels: male and female) was conducted. Further, paired-samples t-tests and post hoc tests using Bonferroni were conducted to assess whether faces with high and low levels of skin smoothness, baby facedness and masculinity differed in the expected ways regarding their trait ratings.

To test whether there was an interaction effect of gender and facial skin smoothness levels (high vs. low) on personality trait ratings (EQ1), a two-way repeated measures ANOVA was conducted with the within-subjects factors *Manipulation Level* (two levels: *High and Low*) and *Gender* (two levels: *Male and Female*) for each personality trait. To test how facial baby facedness affected impressions of Openness,

Extraversion and Neuroticism (EQ2), paired-samples t-tests and post hoc tests using Bonferroni were conducted to assess whether faces with high and low levels of baby facedness differed in the respective personality trait ratings and, if yes, how exactly they differed. Further, to test how facial masculinity affected impressions of Conscientiousness, Extraversion and Neuroticism (EQ3), again, paired-samples t-tests and post hoc tests using Bonferroni were conducted to assess whether faces with high and low levels of masculinity differed in the respective personality trait ratings and, if yes, how exactly they differed. Finally, to test whether there was an interaction effect of gender and facial masculinity levels (high vs. low) on personality trait ratings (EQ4), a two-way repeated measures ANOVA was conducted with the within-subjects factors *Manipulation Level* (two levels: *High and Low*) and *Gender* (two levels: *Male and Female*) for each personality trait.

CHAPTER 5 RESULTS

5. 1. NORMALITY TESTS

The effects of face feature manipulations and gender were assessed separately for each of the eight trait ratings (Psychopathy, Narcissism, Machiavellianism, Neuroticism, Extraversion, Openness, Conscientiousness and Agreeableness). Prior to the main analyses, the normality distribution of each trait rating with respect to face feature (skin smoothness, baby facedness, masculinity), manipulation level (high vs. low) and gender (male vs. female) was assessed. Kolmogorov-Smirnov tests indicated that none of the mean trait ratings followed a normal distribution ($p < .001$). However, when looking at the absolute values of skewness and kurtosis, it was seen that all skewness and kurtosis absolute values of the trait ratings were ≤ 1.10. According to Mishra et a (2019), for sample sizes above 300, an absolute skewness of ≤2 and an absolute kurtosis of ≤4 can be used as reference values for inferring sufficient normality. Thus, the skewness and kurtosis values of the data would suggest the presence of reasonable normality. Further, visual assessment of the histograms for each trait rating revealed approximately bell-shaped distributions for most trait ratings, which also supports the inference that the data is sufficiently normally distributed (Mishra et al., 2019). Finally, the data did not contain any trait ratings with an absolute z-score of ± 3.29. Thus, no outliers were detected, and all participants (N= 505) were included in the confirmatory and exploratory analyses.

5. 2. CONFIRMATORY ANALYSES

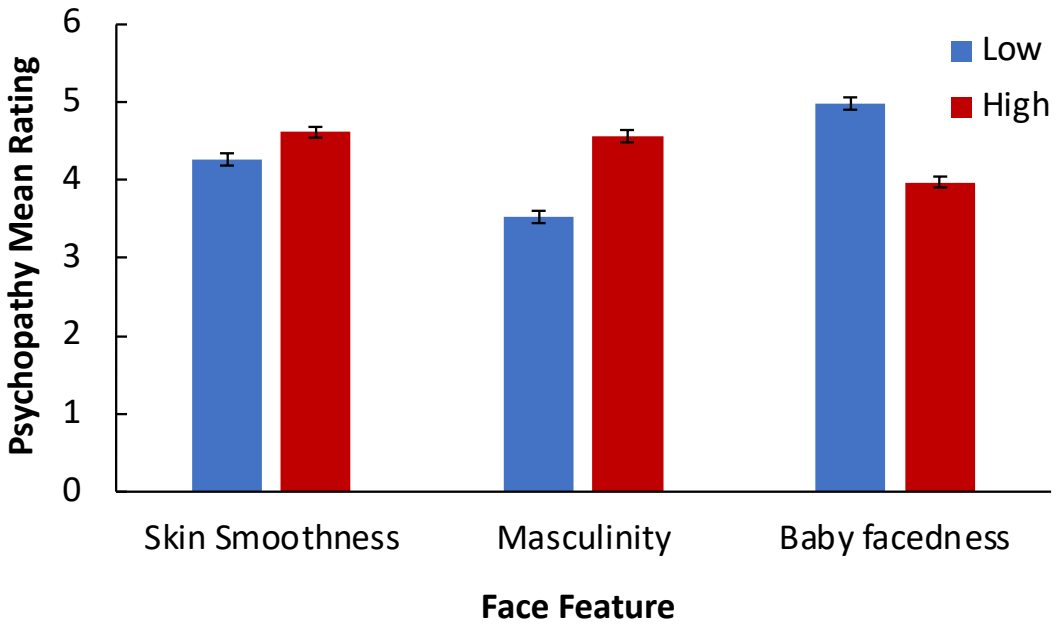
5.2.1. Psychopathy

To assess whether psychopathy ratings differed between the two manipulation levels across our three face feature categories, a three-way repeated measures ANOVA was conducted with the within-subjects factors of *Face Feature* (three levels: *Skin Smoothness, Masculinity and Baby facedness*), *Manipulation Level* (two levels: *High*

and Low) and *Gender* (two levels: *Male and Female*). Sphericity was violated for the interaction between face feature and manipulation level ($p < .01$), thus Greenhouse-Geisser corrected results are reported.

To support hypotheses H1, H2 and H3, there should be a significant interaction between face feature and manipulation levels, such that low level skin smoothness and baby facedness are rated as higher on psychopathy compared to high level skin smoothness and baby facedness while high level of masculinity is rated as higher on psychopathy compared to low level masculinity. Results showed that there was indeed a statistically significant interaction between Face Feature and Manipulation Level, $F(1.96, 989.82) = 156.125$, $p < .001$, $\eta^2 = .237$ (Figure 5.1.). Post hoc tests using Bonferroni corrected paired samples t-test revealed that the effect of baby facedness level (high vs. low) and masculinity level (high vs. low) on psychopathy ratings was as expected: Babyish faces ($M = 4.979$, $SE = .076$) were rated as lower on psychopathy compared to mature faces ($M = 3.969$, $SE = .071$; $MD = -1.01$, $SE = .086$, $p_{bonferroni} < .001$; $t(504) = -11.81$, $p < .001$, 95% C.I. [0.842, 1.178], $d = -0.53$) and masculine faces ($M = 4.561$, $SE = .075$) were rated as higher on psychopathy compared to feminine faces ($M = 3.528$, $SE = .072$; $MD = 1.034$, $SE = .087$, $p_{bonferroni} < .001$; $t(504) = 11.94$, $p < .001$, 95% C.I. [0.864, 1.204], $d = 0.53$). However, the effect of skin smoothness on psychopathy ratings performed in the opposite direction of what was expected (Figure 5.1). Namely, smoothed faces ($M = 4.612$, $SE = .074$) were rated as higher on psychopathy compared to blemished faces ($M = 4.263$, $SE = .076$; $MD = 0.349$, $SE = .074$, $p_{bonferroni} < .001$; $t(504) = 4.75$, $p < .001$, 95% C.I. [0.204, 0.493], $d = .21$).

Thus, H1, the hypothesis that faces with blemished skin (i.e., low level skin smoothness) would be judged as having a higher score Psychopathy compared to faces with smoothed skin (i.e., high level skin smoothness) was not supported as the complete opposite was found. However, results supported both H2 and H3, the hypotheses that mature faces (i.e., low level baby facedness) would be judged as having a higher score on Psychopathy compared to babyish faces (i.e., high level baby facedness) and that masculine faces (i.e., high level masculinity) would be judged as having a higher score on Psychopathy compared to feminine faces (i.e., low level masculinity).



Note. Error bars represent SE Means.

Figure 5.1. Psychopathy Ratings for Faces Low and High on Different Face Features.

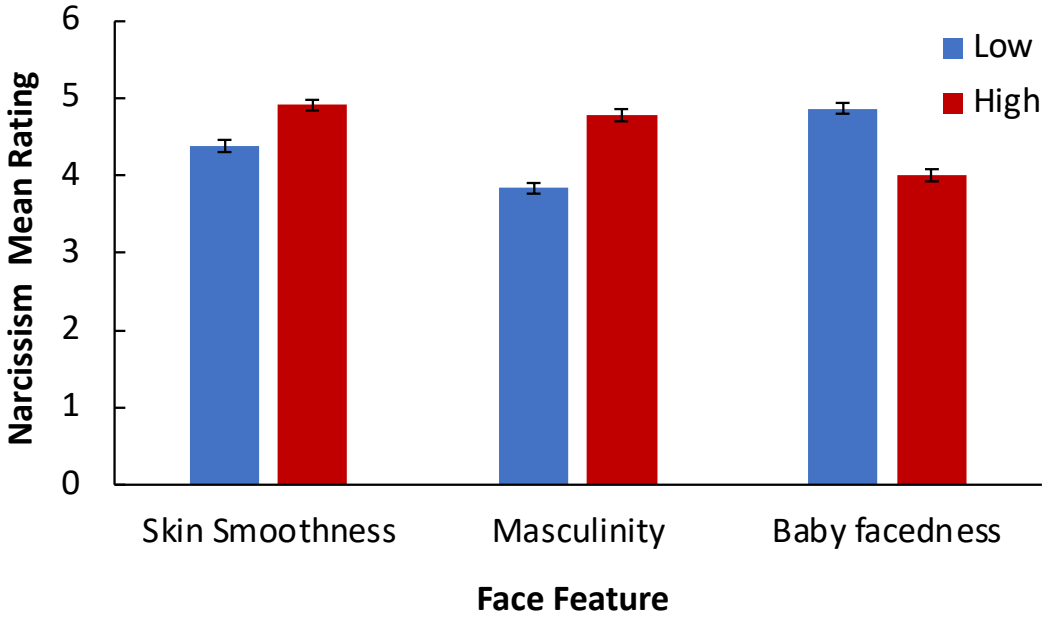
5.2.2. Narcissism

To assess whether Narcissism ratings differed between the two manipulation levels across our three face feature categories, a three-way repeated measures ANOVA was conducted with the within-subjects factors of *Face Feature* (three levels: *Skin Smoothness, Masculinity and Baby facedness*), *Manipulation Level* (two levels: *High and Low*) and *Gender* (two levels: *Male and Female*).

To support hypotheses H1, H2 and H3, there should be a significant interaction between face feature and manipulation levels, such that high level skin smoothness and masculinity are rated as higher on narcissism compared to low level skin smoothness and masculinity while high level baby facedness is rated as lower on narcissism compared to low level baby facedness. Results showed that there was indeed a statistically significant interaction between Face Feature and Manipulation Level, $F(2, 1008) = 122.632$, $p < .001$, $\eta^2 = .196$, and that the effect of skin smoothness (high vs. low), masculinity (high vs. low) and baby facedness (high vs. low) performed as expected. Post hoc tests using Bonferroni corrected paired samples t-test revealed that smoothed faces ($M = 4.910$, $SE = .071$) were rated as higher on narcissism compared to blemished faces ($M = 4.385$, $SE = .074$; $MD = 0.525$, $SE = .074$, $p_{bonferroni}$

< .001; $t(504) = 7.14$, $p < .001$, 95% C.I. [0.380, 0.669]; $d = 0.32$). Masculine faces ($M = 4.778$, $SE = .077$) were rated as higher on narcissism compared to feminine faces ($M = 3.839$, $SE = .070$; $MD = 0.940$, $SE = .091$, $p_{bonferroni} < .001$; $t(504) = 10.3$, $p < .001$, 95% C.I. = [0.760, 1.119], $d = 0.46$) and, finally, babyish faces ($M = 4.004$, $SE = .075$) were rated as lower on narcissism compared to mature faces ($M = 4.868$, $SE = .076$; $MD = -0.864$, $SE = .087$, $p_{bonferroni} < .001$; $t(504) = -9.89$, $p < .001$, 95% C.I. = [-1.036, -0.693], $d = -0.44$; Figure 5.2.)

Thus, H1, the hypothesis that faces with smoothed skin (i.e., high level skin smoothness) would be judged as having a higher score Narcissism compared to faces with blemished skin (i.e., low level skin smoothness) was not supported. Results also supported both H2 and H3, the hypotheses that mature faces (i.e., low level baby facedness) would be judged as having a higher score on Narcissism compared to babyish faces (i.e., high level baby facedness) and that masculine faces (i.e., high level masculinity) would be judged as having a higher score on Narcissism compared to feminine faces (i.e., low level masculinity).



Note. Error bars represent SE Means.

Figure 5.2. Narcissism Ratings for Faces Low and High on Different Face Features.

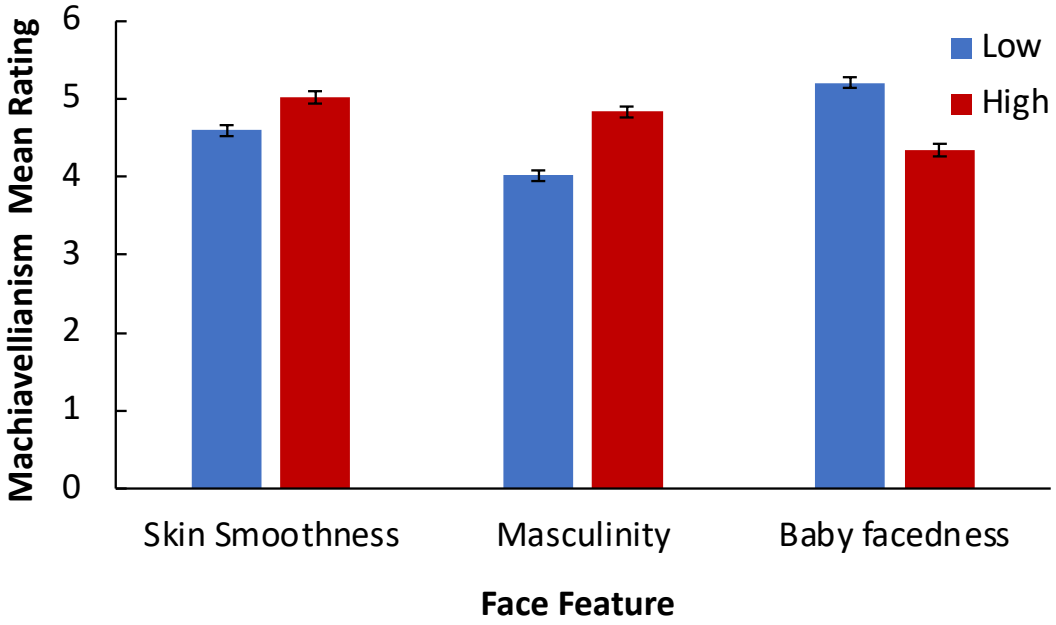
5.2.3. Machiavellianism

To assess whether Machiavellianism ratings differed between the two manipulation levels across our three face feature categories, a three-way repeated measures ANOVA was conducted with the within-subjects factors of *Face Feature* (three levels: *Skin Smoothness, Masculinity and Baby facedness*), *Manipulation Level* (two levels: *High and Low*) and *Gender* (two levels: *Male and Female*). Sphericity was violated for the interaction between face and manipulation level ($p < .001$), thus, Greenhouse-Geisser corrected results are reported.

To support hypotheses H1, H2 and H3, there should be a significant interaction between face feature and manipulation levels, such that low level skin smoothness and baby facedness are rated as higher on Machiavellianism compared to high level skin smoothness and baby facedness while high level masculinity is rated as higher on Machiavellianism compared to low level masculinity. Results showed that there was indeed a statistically significant interaction between Face Feature and Manipulation Level, $F(1.92,968.09) = 104.314$, $p < .001$, $\eta^2 = .171$ (Figure 5.3.). Post hoc tests using Bonferroni corrected paired samples t-test revealed that the effect of baby facedness level (high vs. low) and masculinity level (high vs. low) on Machiavellianism ratings was as expected: Babyish faces ($M = 4.339$, $SE = .073$) were rated as lower on Machiavellianism compared to mature faces ($M = 5.204$, $SE = .075$; $MD = -0.865$, $SE = .088$, $p_{bonferroni} < .001$; $t(504) = -9.88$, $p < .001$, 95% C.I. = [0.693, 1.038], $d = -0.44$) and masculine faces ($M = 4.831$, $SE = .076$) were rated as higher on Machiavellianism compared to feminine faces ($M = 4.014$, $SE = .073$; $MD = 0.817$, $SE = .089$, $p_{bonferroni} < .001$; $t(504) = 9.14$, $p < .001$, 95% C.I. = [0.641, 0.992], $d = 0.41$). However, the effect of skin smoothness on Machiavellianism ratings performed in the opposite direction of what was expected (Figure 5.3.). Namely, smoothed faces ($M = 5.021$, $SE = .075$) were rated as higher on Machiavellianism compared to blemished faces ($M = 4.592$, $SE = .076$; $MD = 0.429$, $SE = .076$, $p_{bonferroni} < .001$; $t(504) = 5.63$, $p < .001$, 95% C.I. = [0.279, 0.578], $d. = 0.25$).

Thus, H1, the hypothesis that faces with blemished skin (i.e., low level skin smoothness) would be judged as having a higher score Machiavellianism compared to faces with smoothed skin (i.e., high level skin smoothness) was not supported since the complete opposite was found. However, results supported both H2 and H3, the hypotheses that mature faces (i.e., low level baby facedness) would be judged as

having a higher score on Machiavellianism compared to babyish faces (i.e., high level baby facedness) and that masculine faces (i.e., high level masculinity) would be judged as having a higher score on Machiavellianism compared to feminine faces (i.e., low level masculinity).



Nore. Error bars represent SE Means.

Figure 5.3. Machiavellianism Ratings for Faces Low and High on Different Face Features

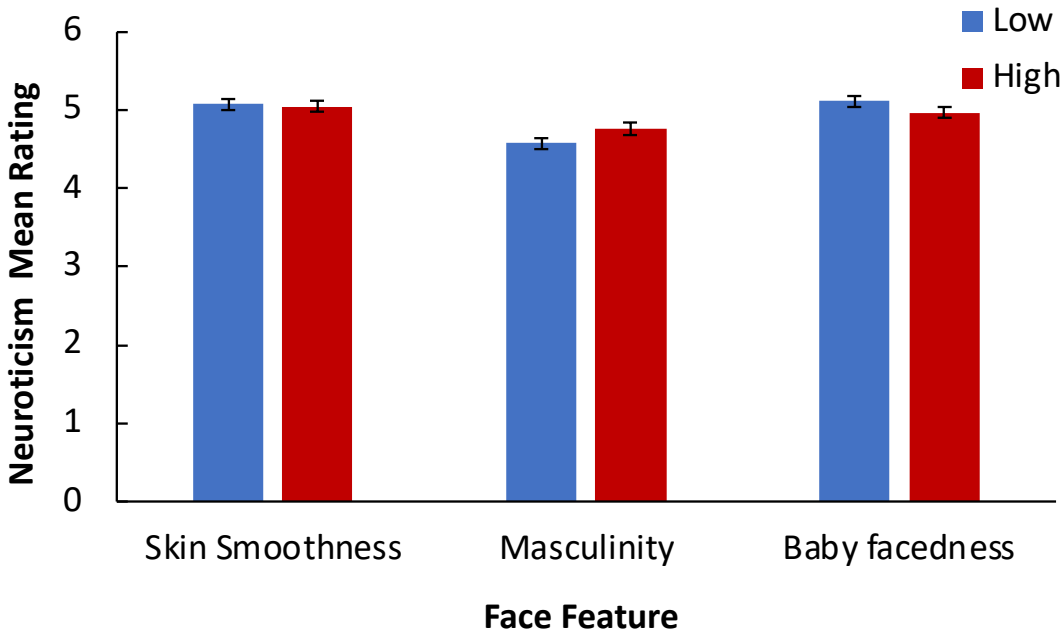
5.2.4. Neuroticism

To assess whether Neuroticism ratings differed between the two manipulation levels of skin smoothness, a three-way repeated measures ANOVA was conducted with the within-subjects factors of *Face Feature* (three levels: *Skin Smoothness, Masculinity and Baby facedness*), *Manipulation Level* (two levels: *High and Low*) and *Gender* (two levels: *Male and Female*). Sphericity was violated for the interaction between face and manipulation level ($p < .001$), thus Greenhouse-Geisser corrected results are reported.

To support hypothesis H1, there should be a significant interaction between face feature and manipulation levels, such that low level skin smoothness is rated as higher on neuroticism compared to high level skin smoothness. Results showed that there was indeed a statistically significant interaction between Face Feature and Manipulation Level, $F(1.93,969.54) = 49.769$, $p < .05$, $\eta^2 = .008$). However, post hoc tests using Bonferroni corrected paired samples t-test revealed that faces with high (*M*

= 5.041, *SE* = .069) and low level (*M* = 5.069, SE = .073) skin smoothness did not appear to significantly differ in their neuroticism ratings (*MD* = -0.029, *SE* = .070, $p_{bonferroni}$ = 1; *t*(504) = -0.41, *p* = .681, 95% C.I. [-0.17, 0.10], *d* = -0.02; Figure 5.4.).

Thus, H1, the hypothesis that faces with blemished skin (i.e., low level skin smoothness) would be judged as having a higher score neuroticism compared to faces with smoothed skin (i.e., high level skin smoothness), was not supported since neuroticism ratings did not appear to differ based on manipulation levels.



Note. Error bars represent SE Means.

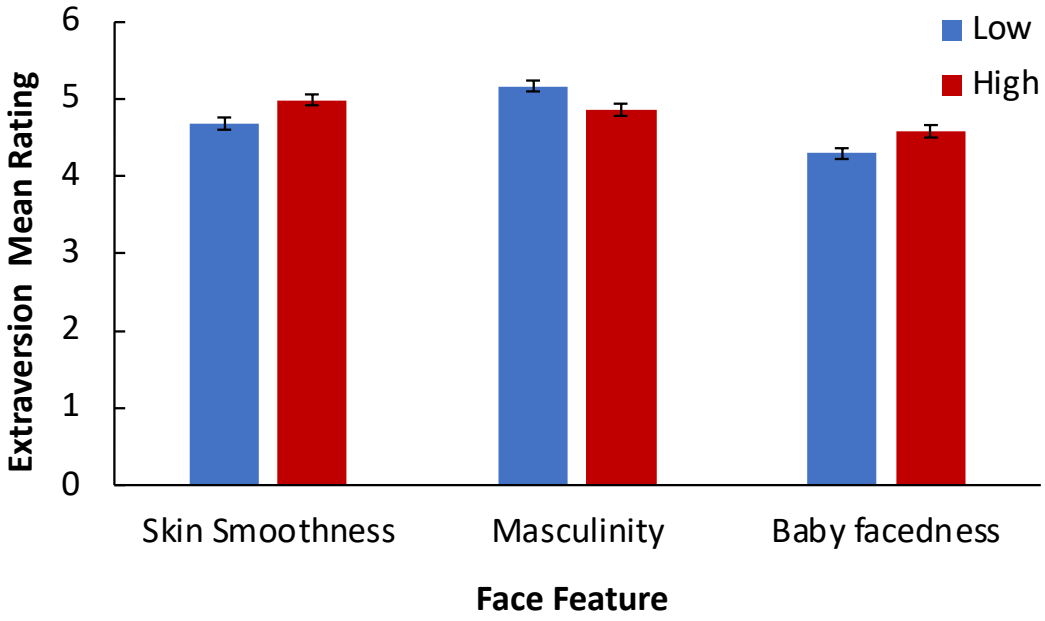
Figure 5.4. Neuroticism Ratings for Faces Low and High on Different Face Features

5.2.5. Extraversion

To assess whether extraversion ratings differed between the two manipulation levels of skin smoothness, a three-way repeated measures ANOVA was conducted with the within-subjects factors of *Face Feature* (three levels: *Skin Smoothness, Masculinity and Baby facedness*), *Manipulation Level* (two levels: *High and Low*) and *Gender* (two levels: *Male and Female*).

To support hypothesis H1, there should be a significant interaction between face feature and manipulation levels, such that high level skin smoothness is rated as higher on extraversion compared to low level skin smoothness. Results showed that there was indeed a statistically significant interaction between Face Feature and Manipulation Level, *F*(2, 1008) = 20.495, *p* < .001, η^2 = .039 (Figure 5.5.). Namely,

post hoc tests using Bonferroni corrected paired samples t-test revealed that, smoothed skin faces (M = 4.981, SE = .068) were rated as significantly higher on extraversion compared to blemished faces (M = 4.680, SE = .071; MD = 0.301, SE = .073, $p_{bonferroni}$ < .001; $t(504) = 4.14$, $p < .001$, 95% C.I. = [0.158, 0.444], d = 0.18). Thus, results were in favour of H1.



Note. Error bars represent SE Means.

Figure 5.5. Extraversion Ratings for Faces Low and High on Different Face Features

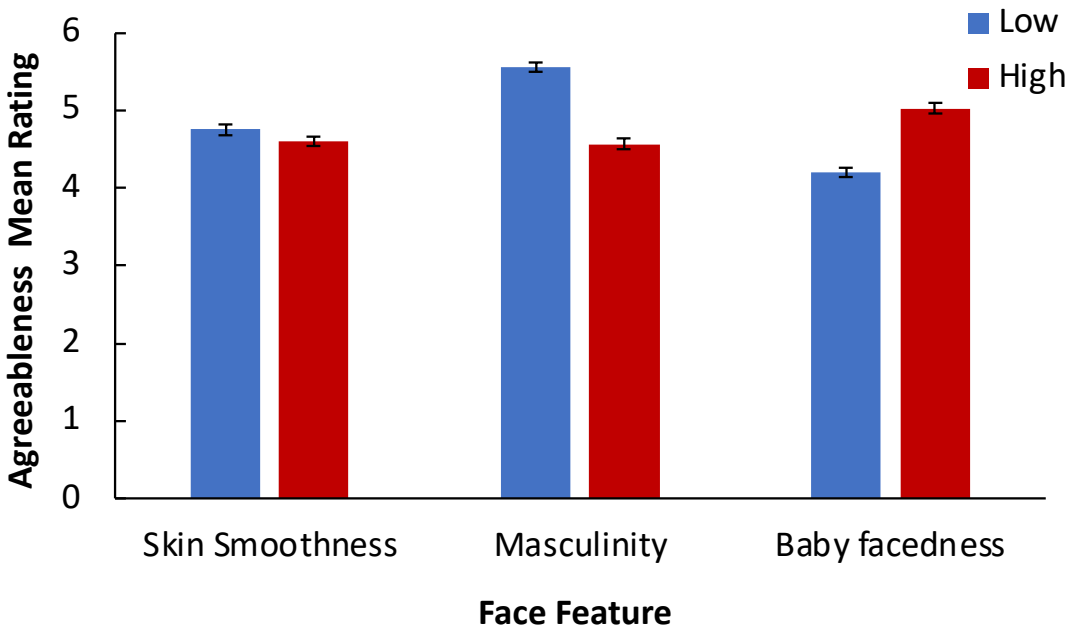
5.2.6. Agreeableness

To assess whether agreeableness ratings differed between the two manipulation levels across our three face feature categories, a three-way repeated measures ANOVA was conducted with the within-subjects factors of *Face Feature* (three levels: *Skin Smoothness, Masculinity and Baby facedness*), *Manipulation Level* (two levels: *High and Low*) and *Gender* (two levels: *Male and Female*).

To support hypotheses H1, H2 and H3, there should be a significant interaction between face feature and manipulation levels, such that high level skin smoothness and baby facedness are rated as higher on agreeableness compared to low level skin smoothness and baby facedness while low level masculinity is rated as higher on agreeableness compared to high level masculinity. Results showed that there was indeed a statistically significant interaction between Face Feature and Manipulation Level, $F(2, 1008) = 125.894$, $p < .001$, $\eta^2 = .200$ (Figure 5.6.). Post hoc tests using

Bonferroni corrected paired samples t-test revealed that the effects of baby facedness level (high vs. low) and masculinity level (high vs. low) on agreeableness ratings was as expected, babyish faces (M = 5.022, SE = .066) were rated as higher on agreeableness compared to mature faces (M = 4.204, SE = .066; MD = 0.818, SE = .085, $p_{bonferroni} < .001$; $t(504) = 9.59$, $p < .001$, 95% C.I. = [0.650, 0.985], d = 0.43), and masculine faces (M = 4.566, SE = .066) were rated as lower on agreeableness compared to feminine faces (M = 5.558, SE = .064; MD = -0.992, SE = .080, $p_{bonferroni} < .001$; $t(504) = -12.39$, $p < .001$, 95% C.I. = [0.835, 1.149], d = -0.55). For skin smoothness, Bonferroni corrected paired samples t-test showed that there was no significant difference in agreeableness ratings between the two manipulation levels of skin smoothness (MD = -0.151, SE = .072, $p_{bonferroni}$ = .568). However, without Bonferroni correction, smoothed skin faces (M = 4.602, SE = 065) were rated as slightly lower on agreeableness compared to blemished faces (M = 4.752, SE = .068; $t(504) = -2.08$, $p < .05$, 95% C.I. = [0.008, 0.293], d = -0.09).

Thus, H1, the hypothesis that faces with smoothed skin (i.e., high level skin smoothness) would be judged as having a higher score agreeableness compared to faces with blemished skin (i.e., low level skin smoothness), was not supported since levels of skin smoothness did not affect agreeableness ratings. However, findings showed support for both H2 and H3, namely the hypotheses that babyish faces (i.e., high level baby facedness) would be judged as having a higher score on agreeableness compared to mature faces (i.e., low level baby facedness) and that feminine faces (i.e., low level masculinity) would be judged as having a higher score on agreeableness compared to masculine faces (i.e., high level masculinity).



Note. Error bars represent SE Means.

Figure 5.6. Agreeableness Ratings for Faces Low and High on Different Face Features

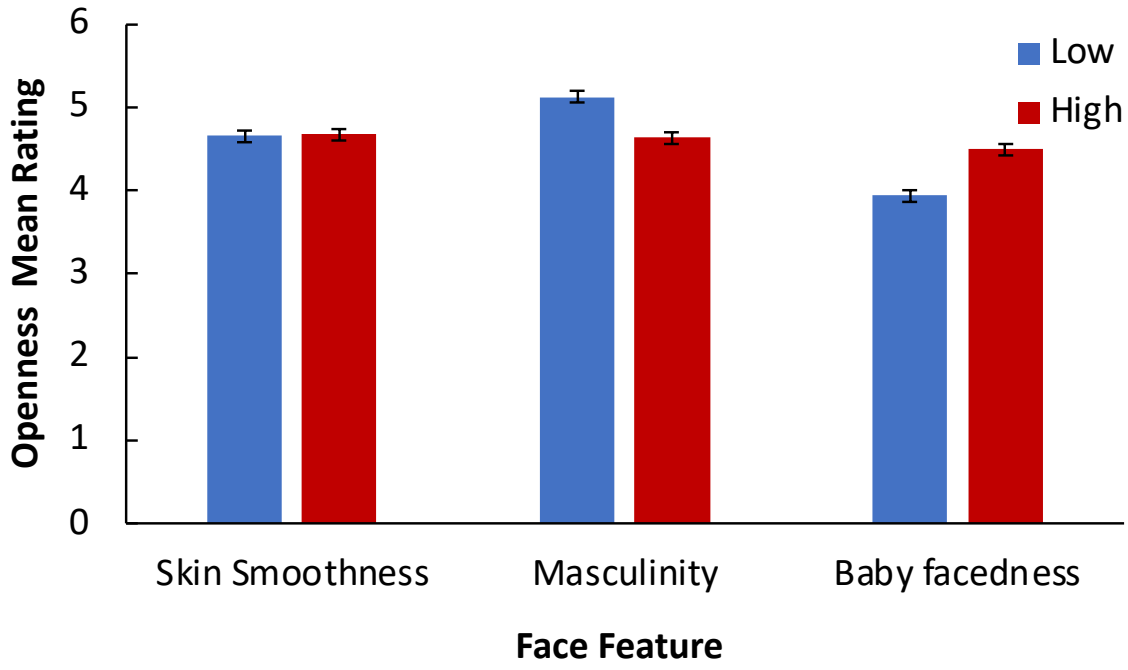
5.2.7. Openness

To assess whether openness ratings did differ between the two manipulation levels across the two face feature categories skin smoothness and masculinity, a three-way repeated measures ANOVA was conducted with the within-subjects factors of *Face Feature* (three levels: *Skin Smoothness, Masculinity and Baby facedness*), *Manipulation Level* (two levels: *High and Low*) and *Gender* (two levels: *Male and Female*). Sphericity was violated for the interaction between face feature and manipulation level ($p < .045$), thus Greenhouse-Geisser corrected results are reported.

To support hypotheses H1 and H3, there should be a significant interaction between face feature and manipulation levels, such that high level skin smoothness is rated as higher on openness compared to low level skin smoothness while low level masculinity is rated as higher on openness compared to high level masculinity. Results showed that there was indeed a statistically significant interaction between Face Feature and Manipulation Level, $F(1.98,995.77) = 49.769$, $p < .001$, $\eta^2 = .090$ (Figure 5.7). Post hoc tests using Bonferroni corrected paired samples t-test revealed that the effect of masculinity level (high vs. low) on openness ratings was as expected, feminine faces ($M = 5.127$, $SE = .068$) were rated as higher on openness compared to

masculine faces (M = 5.127, SE = .068; MD = -0.493, SE = .081, $p_{bonferroni}$ < .001; t(504) = -6.35, p < .001, 95% C.I. = [0.341, 0.646], d = -0.28). However, faces with high (M = 4.672, SE = .067) and low (M= 4.654, SE = .069) levels of skin smoothness did not appear to differ in their openness ratings (MD = 0.018, SE = .065, $p_{bonferroni}$ = 1; t(504) = .274, p = .784, 95% C.I. [-0.11, 0.15], d = 0.01).

Thus, H1, the hypothesis that faces with high levels of skin smoothness would be rated as higher on openness compared to faces with low levels of skin smoothness, was not supported since openness ratings of skin smoothness faces did not appear to differ based on manipulation levels. However, the results showed support for H3, the hypothesis that feminine faces (i.e., low level masculinity) would be judged as having a higher score on openness compared to masculine faces (i.e., high level masculinity).



Note. Error bars represent SE Means.

Figure 5.7. Openness Ratings for Faces Low and High on Different Face Features.

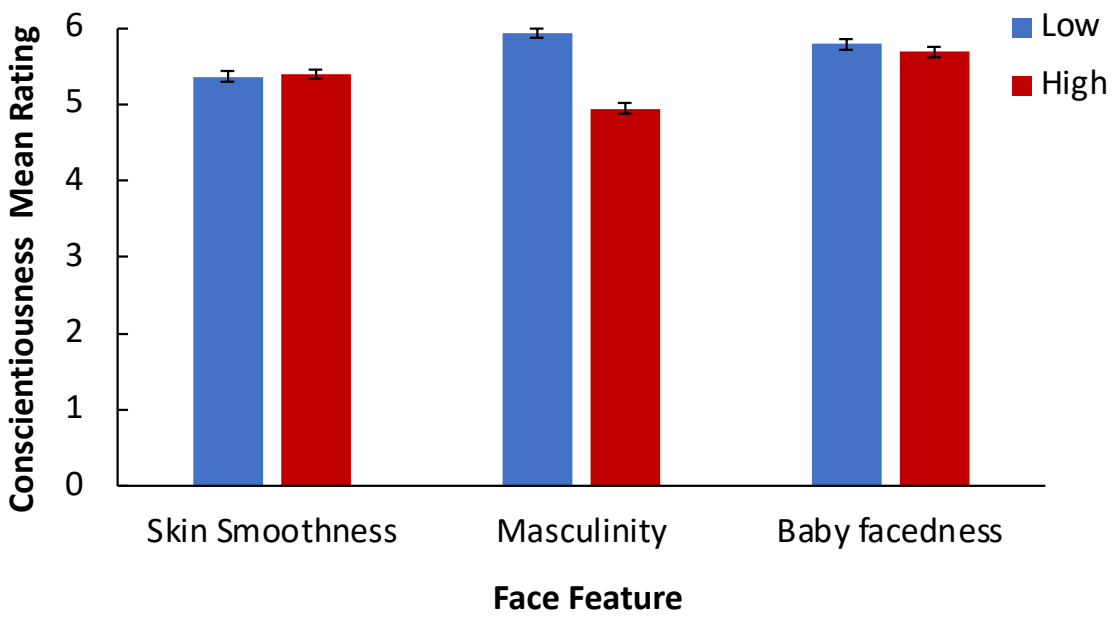
5.2.8. Conscientiousness

To assess whether conscientiousness ratings differed between the two manipulation levels across the two face feature categories skin smoothness and baby facedness, a three-way repeated measures ANOVA was conducted with the within-subjects factors of *Face Feature* (three levels: *Skin Smoothness, Masculinity and Baby*

facedness), *Manipulation Level* (two levels: *High and Low*) and *Gender* (two levels: *Male and Female*).

To support hypotheses H1 and H2 there should be a significant interaction between face feature and manipulation levels, such that high level skin smoothness is rated as higher on conscientiousness compared to low level skin smoothness while low level baby facedness is rated as higher on conscientiousness compared to high level baby facedness. Results showed that there was indeed a statistically significant interaction between Face Feature and Manipulation Level, $F(2, 1008) = 54.953$, $p < .001$, $\eta^2 = .098$ (Figure 5.8.). However, post hoc tests using Bonferroni corrected paired samples t-test revealed that, faces with high ($M = 5.396$, $SE = .067$) and low ($M = 5.360$, $SE = .067$) level skin smoothness did not significantly differ in their conscientiousness ratings ($MD = 0.036$, $SE = .067$, $p_{bonferroni} = 1$; $t(504) = 0.532$, $p = .595$, 95% C.I. [-0.1, 0.17], $d = 0.02$). Further, faces with high ($M = 5.691$, $SE = .068$) and low ($M = 5.785$, $SE = .068$) level baby facedness also did not appear to notably differ in their conscientiousness ratings ($MD = -0.094$, $SE = .079$, $p_{bonferroni} = 1$; $t(504) = -1.19$, $p = .079$, 95% C.I. [-0.25, 0.06], $d = -0.05$). Rather, this interaction appeared to be driven by differences in high and low level facial masculinity (Figure 5.8.).

Thus, H1, the hypothesis that faces with smoothed skin (i.e., high level skin smoothness) would be judged as having a higher score conscientiousness compared to faces with blemished skin (i.e., low level skin smoothness), was not supported since conscientiousness ratings of skin smoothness faces did not seem to differ based on manipulation levels. Similarly, H2, the hypothesis that mature faces (i.e., low level baby facedness) would be judged as having a higher score on conscientiousness compared to babyish faces (i.e., high level baby facedness), was also not supported since, again, conscientiousness ratings did not seem to differ based on baby facedness manipulation levels.



Note. Error bars represent SE Means.

Figure 5.8. Conscientiousness Ratings for Faces Low and High on Different Face Features.

5. 3. EXPLORATORY ANALYSES

5. 3. 1. Gender and Skin Smoothness Interaction

To assess whether there was an interaction effect of gender and facial skin smoothness levels (high vs. low) on personality trait ratings (EQ1), a two-way repeated measures ANOVA was conducted with the within-subjects factors *Manipulation Level* (two levels: *High and Low*) and *Gender* (two levels: *Male and Female*) for each personality trait.

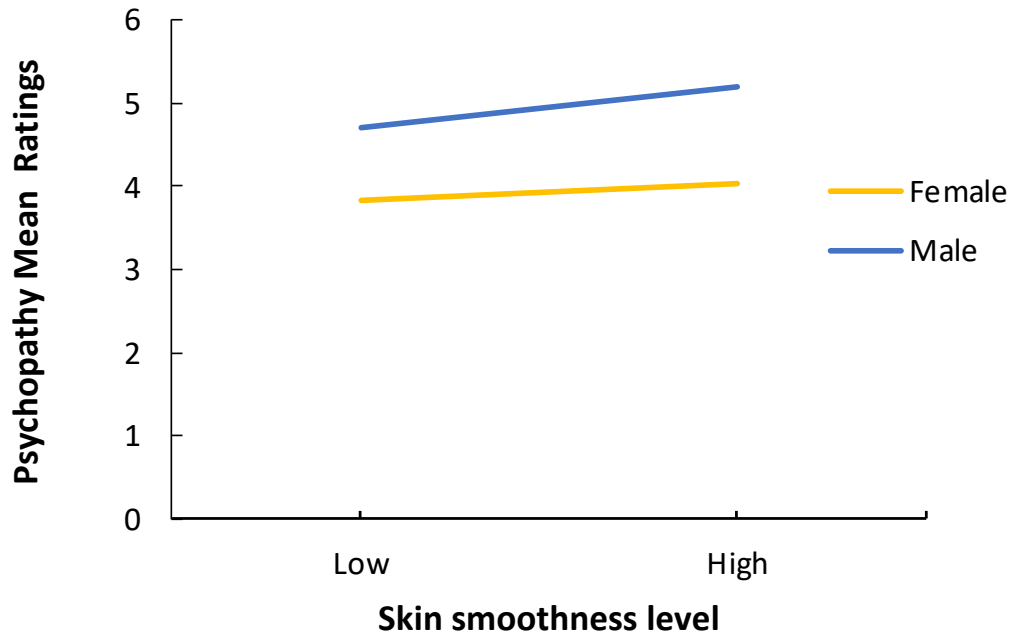


Figure 5.9. Interaction Between Skin Smoothness Level and Gender on Psychopathy Ratings

For Psychopathy, results showed that there was a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 5.417$, $p < .05$, $\eta^2 = .01$. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for male faces, smoothed faces ($M = 5.2$, $SE = .1$) were rated as higher on psychopathy compared to blemished faces ($M = 4.7$, $SE = .1$; $MD = 0.497$, $SE = .10$, $p_{bonferroni} < .001$, $t(504) = 4.91$, $p < .001$, $d = 0.22$). However, for female faces, smoothed ($M = 4.02$, $SE = .1$) and blemished faces ($M = 3.82$, $SE = .1$) did not appear to significantly differ in their psychopathy ratings ($MD = -1.176$, $SE = .13$, $p_{bonferroni} = 1$), though they did seem to slightly differ when Bonferroni corrections were omitted, $t(504) = 2.15$, $p < .05$, $d = 0.10$ (Figure 5.9.).

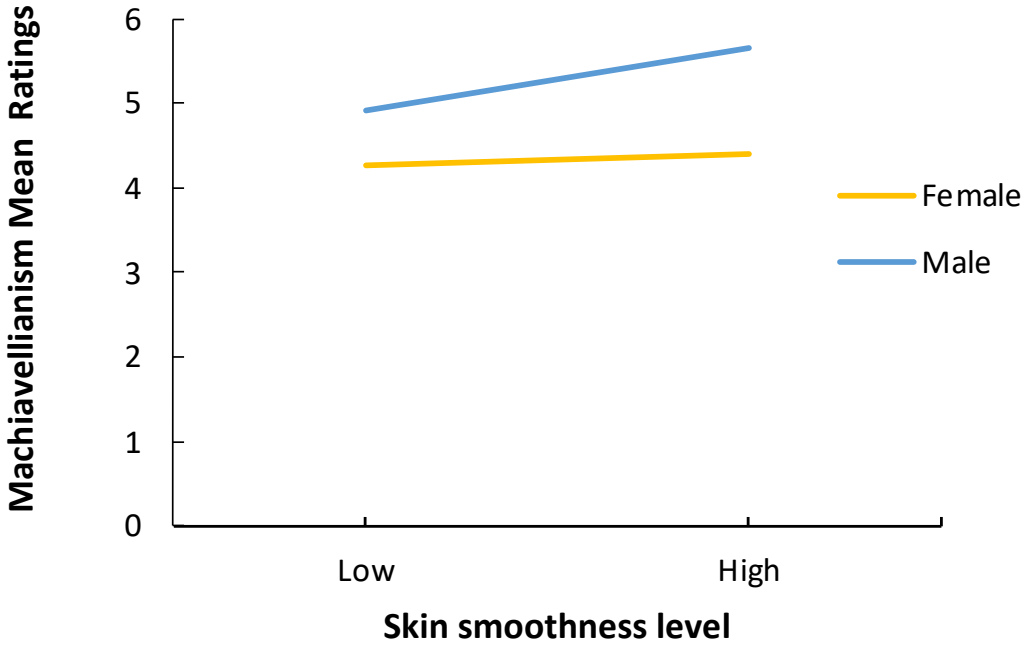


Figure 5.10. Interaction Between Skin Smoothness Level and Gender on Machiavellianism Ratings

For Machiavellianism, results showed that there was indeed a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 19.493$, $p < .001$, $\eta^2 = .04$. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for female faces, smoothed ($M = 4.4$, $SE = .1$) and blemished faces ($M = 4.3$, $SE = .1$) did not appear to notably differ in their Machiavellianism ratings ($MD = 0.127$, $SE = .10$, $p_{bonferroni} = 1$; $t(504) = 1.27$, $p = .203$, $d = 0.06$), whereas for males, smoothed skin faces ($M = 5.6$, $SE = .1$) were rated as higher on Machiavellianism compared to blemished faces (M = 4.92, SE = .096; $MD = 0.731$, $SE = .11$, $p_{bonferroni} < .001$; $t(504) = 6.94$, $p < .001$, $d = 0.31$; Figure 5.10.).

Figure 5.11. Interaction Between Skin Smoothness Level and Gender on Narcissism Ratings

For Narcissism ratings, results showed that there was a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 44.858$, $p < .001$, $\eta^2 = .08$. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for female faces, smoothed ($M = 3.82$, $SE = .1$) and blemished faces ($M = 3.74$, $SE = .09$) did not appear to notably differ in their Narcissism ratings ($MD = 0.077$, $SE = .10$, $p_{bonferroni} = 1$; $t(504) = .796$, $p = .426$, $d = 0.04$, whereas for males, smoothed skin faces (M = 6.00, SE = .1) were rated as higher on Narcissism compared to blemished faces ($M = 5.03$, $SE = .11$; $MD = 0.972$, $SE = .10$, $p_{bonferroni} < .001$; $t(504) = 9.57$, $p < .001$, $d = 0.43$; Figure 5.11.).

Figure 5.12. Interaction Between Skin Smoothness Level and Gender on Extraversion Ratings

For Extraversion, results showed that there was indeed a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 4.387$, $p < .05$, $\eta^2 = .01$. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for female faces, smoothed skin faces ($M = 4.65$, $SE = .09$) and blemished faces ($M = 4.48$, $SE = .09$) did not significantly differ in Extraversion ratings ($MD = 0.168$, $SE = .09$, $p_{bonferroni} = 1$; $t(504) = 1.84$, $p = .07$, $d = 0.08$), whereas for males, smoothed skin faces ($M = 5.31$, $SE = .09$) were rated as higher on Extraversion compared to blemished faces ($M = 4.88$, $SE = .1$; $MD = 0.434$, $SE = .10$, $p_{bonferroni} = .001 t(504) = 4.29$, $p < .001$, $d = 0.19$; Figure 5.12.).

Figure 5.13. Interaction Between Skin Smoothness Level and Gender on Agreeableness Ratings

For Agreeableness ratings, results showed that there was a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 6.234$, $p < .05$, $\eta^2 = .01$. However, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for female faces, smoothed ($M = 4.96$, $SE = .09$) and blemished faces ($M = 4.96$, $SE = .09$) did not differ in their Agreeableness ratings ($MD = 0.010$, $SE = .10$, $p_{bonferroni} = 1$; $t(504) = 0.60$, $p = .952$, $d = 0.00$). For males too, blemished ($M = 4.24$, $SE = .09$) and smoothed faces ($M = 4.55$, $SE = .09$) did not differ in their Agreeableness ratings ($MD = 0.307$, $SE = .09$, $p_{bonferroni} = .065$), though omitting the Bonferroni correction did reveal a significant difference between the two manipulation levels, $t(504) = -3.32$, $p < .01$, $d = -0.15$ (Figure 5.13.).

Finally, results revealed no statistically significant interaction between Manipulation Level and Gender for Openness ratings, $F(1, 504) = 0.036$, $p = .85$, Neuroticism ratings, $F(1, 504) = 0.728$, $p = .394$, and Conscientiousness ratings, $F(1, 504) = 0.015$, $p = .904$. In other words, male and female faces were rated similarly on Openness, Neuroticism and Conscientiousness in response to the skin smoothness manipulations.

5. 3. 2. Baby Facedness and Impressions of Openness, Extraversion & Neuroticism

Due to the lack of any expectations and hypotheses regarding the effect of baby facedness on Openness, Extraversion and Neuroticism, the corresponding results are reported here as exploratory analyses instead of being reported in the confirmatory analyses.

To test how facial baby facedness affected impressions of Openness, Extraversion and Neuroticism (EQ2), paired-samples t-tests and post hoc tests using Bonferroni were conducted. Post hoc tests using Bonferroni corrected paired samples t-test revealed that, for openness impressions, babyish faces (M = 4.492, SE = .069) were rated as significantly higher on openness compared to mature faces (M = 3.939, SE = .066; MD = 0.554, SE = .079, $p_{bonferroni}$ < .001; $t(504) = 6.98$, p < .001, 95% C.I. = [0.398, 0.709], d = 0.31). For extraversion, post hoc tests using Bonferroni corrected paired samples t-test revealed that, babyish faces (M = 4.584, SE = .070) were rated as significantly higher on extraversion compared to mature faces (M = 4.292, SE = .069; MD = 0.292, SE = .083, $p_{bonferroni}$ = .007; $t(504) = 3.54$, p < .001, 95% C.I. = [0.130, 0.454], d = 0.16). Finally, for neuroticism, post hoc tests using Bonferroni corrected paired samples t-test revealed that, faces with high (M = 4.967, SE = .067) and low (M = 5.108, SE = .068) level baby facedness did not appear to notably differ in their neuroticism ratings (MD = -0.141, SE = .077, $p_{bonferroni}$ = 1; $t(504) =$. -1.82, p = .069, 95% C.I. [-0.29, 0.01], d = -0.08).

5. 3. 3. Masculinity and Impressions of Neuroticism, Extraversion & Conscientiousness

Again, due to the lack of any expectations and hypotheses regarding the effect of masculinity on Neuroticism, Extraversion and Conscientiousness, the corresponding results are reported here as exploratory analyses instead of being reported in the confirmatory analyses.

To test how facial masculinity affected impressions of Neuroticism, Extraversion and Conscientiousness (EQ3), another set of post hoc tests using Bonferroni corrected paired samples t-test revealed that were conducted. Findings showed that masculine faces (M = 4.863, SE = .069) were rated as significantly lower on extraversion compared to feminine faces (M = 5.163, SE = .067; MD = -0.3, SE =

.075, $p_{bonferroni}$ = .001: $t(504) = -4$, $p < .001$, 95% C.I. = [0.153, 0.447], $d = -0.18$). For conscientiousness, masculine faces ($M = 4.941$, $SE = .068$) were rated as significantly lower on conscientiousness compared to feminine faces ($M = 5.931$, $SE = .064$; $MD = -0.990$, $SE = .075$, $p_{bonferroni} < .001$; $t(504) = -13.15$, $p < .001$, 95% C.I. = [0.842, 1.138], $d = -0.59$). However, for neuroticism, masculine ($M = 4.758$, $SE = .072$) and feminine faces ($M = 4.574$, $SE = .070$) did not appear to differ in response to levels of masculinity ($MD = 0.184$, $SE = .086$, $p_{bonferroni} = .463$), even though omitting the Bonferroni correction resulted in a small effect, $t(504) = 2.17$, $p < .05$, 95% C.I. = [0.017, 0.351], $d = 0.1$.

5. 3. 4. Gender and Masculinity Interaction

To test whether there was an interaction effect of gender and facial masculinity levels (high vs. low) on personality trait ratings (EQ4), a two-way repeated measures ANOVA was conducted with the within-subjects factors *Manipulation Level* (two levels: *High and Low*) and *Gender* (two levels: *Male and Female*) for each personality trait.

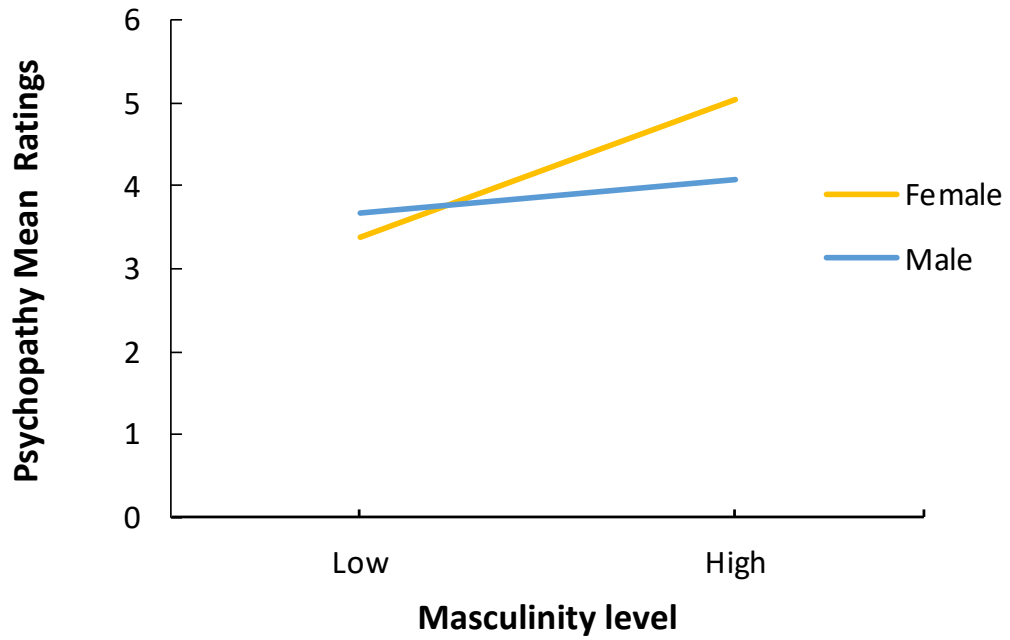


Figure 5.14. Interaction Between Masculinity Level and Gender on Psychopathy Ratings

For psychopathy, results showed that there was indeed a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 64.324$, $p < .001$, $\eta^2 = .11$. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for both male and female faces, masculine faces ($M_{Female} = 5.04$, $SE_{Female} = .1$; $M_{Male} = 4.08$, $SE_{Male} = .1$) were rated as higher on psychopathy compared to feminine faces ($M_{Female} = 3.38$, $SE_{Female} = .09$; $M_{Male} = 3.67$, $SE_{Male} = .1$). However, this difference in psychopathy rating based on masculinity was more prominent for female faces ($MD = 1.657$, $SE = .12$, $p_{bonferroni} < .001$; $t(504) = 13.66$, $p < .001$, $d = 0.61$) compared to male faces($MD = 0.410$, $SE = .11$, $p_{bonferroni} = .017$; $t(504) = 3.69$, $p < .001$, $d = 0.16$; Figure 5.14.).

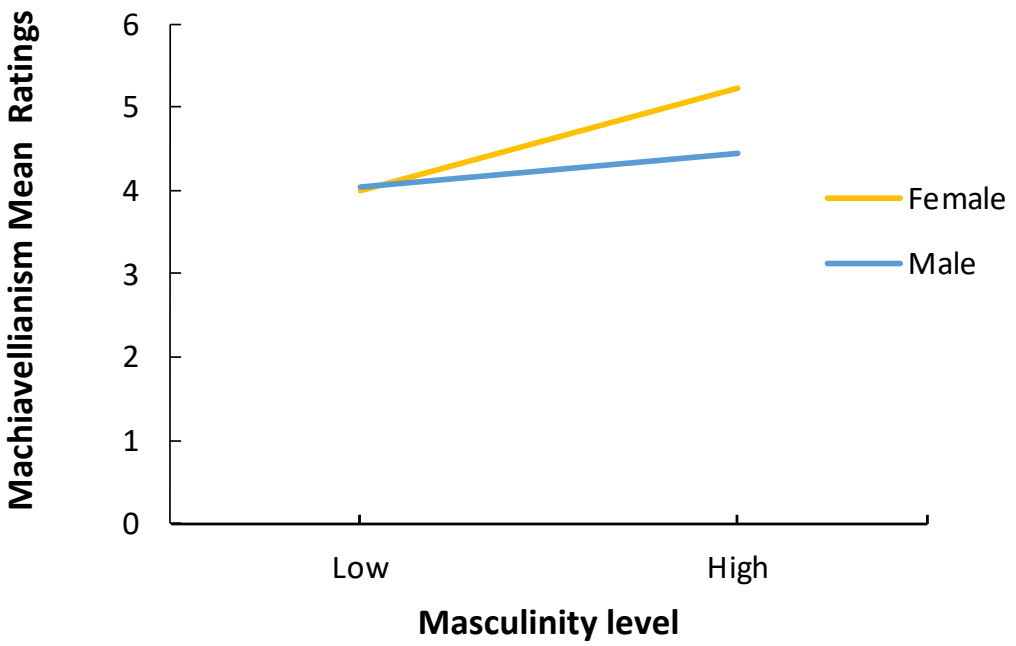


Figure 5.15. Interaction Between Masculinity Level and Gender on Machiavellianism Ratings

For Machiavellianism, results showed that there was indeed a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 28.747$, $p < .001$, $\eta^2 = .05$. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for both male and female faces, masculine faces ($M_{Female} = 5.23$, $SE_{Female} = .1$; $M_{Male} = 4.44$, $SE_{Male} = .1$) were rated as higher on Machiavellianism compared to feminine faces ($M_{Female} = 3.99$, $SE_{Female} = .1$; $M_{Male} = 4.04$, $SE_{Male} = .1$).

However, similar to psychopathy, this difference in Machiavellianism rating based on masculinity was more prominent for female faces (MD = 1.24, SE = .124, $p_{bonferroni}$ < .001; $t(504)$ = 10.01, p < .001, d = 0.45) compared to male faces (MD = 0.394, SE = .114, $p_{bonferroni}$ = .04; $t(504)$ = 3.45, p < .01, d = 0.15; Figure 5.15.).

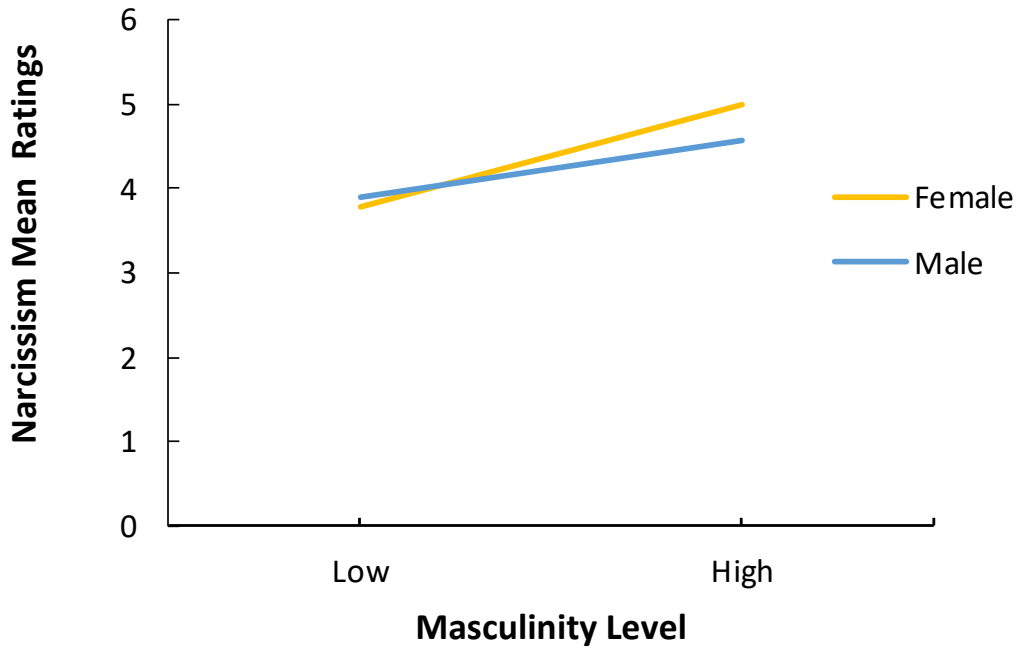


Figure 5.16. Interaction Between Masculinity Level and Gender on Narcissism Ratings

For Narcissism, results showed that there was indeed a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 11.016$, p < .01, η^2 = .02. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for both male and female faces, masculine faces (M_{Female} = 4.99, SE_{Female} = .11; M_{Male} = 4.57, SE_{Male} = .1) were rated as higher on Narcissism compared to feminine faces (M_{Female} = 3.78, SE_{Female} = .09; M_{Male} = 3.9, SE_{Male} = .09). However, similar to psychopathy and Machiavellianism, this difference in Narcissism rating based on masculinity was more prominent for female faces (MD = 1.208, SE = .133, $p_{bonferroni}$ < .001; $t(504)$ = 9.10, p < .001, d = 0.41) compared to male faces (MD = 0.671, SE = .11, $p_{bonferroni}$ < .001; $t(504)$ = 6.10, p < .001, d = 0.27; Figure 5.16.).

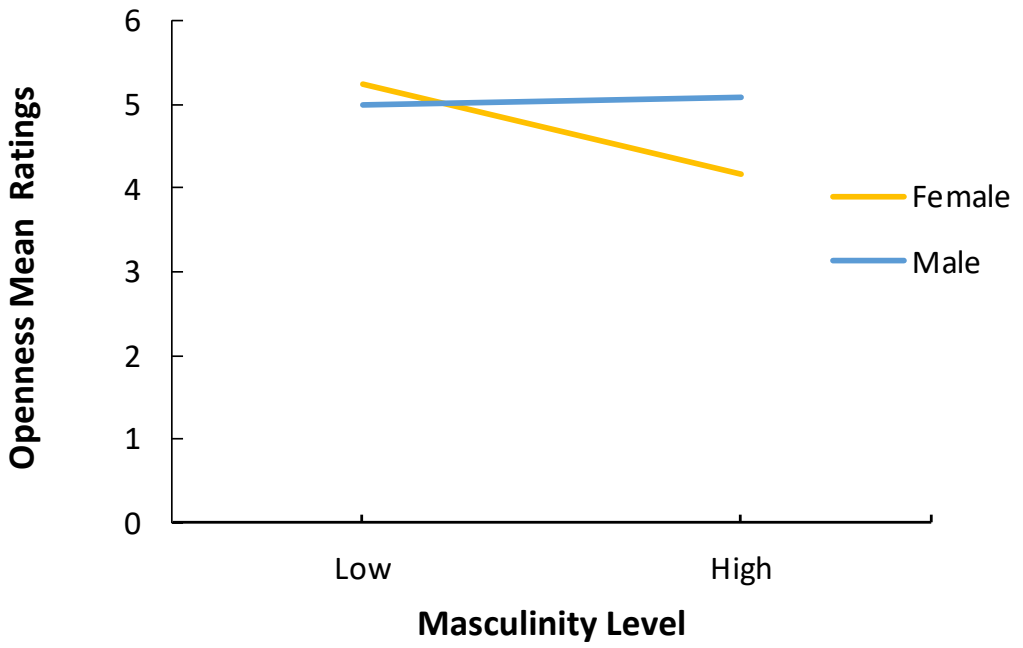


Figure 5.17. Interaction Between Masculinity Level and Gender on Openness Ratings

For Openness, results showed that there was a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 66.904$, $p < .001$, $\eta^2 = .12$. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for male faces, masculine ($M = 5.09$, $SE = .09$) and feminine faces ($M = 5$, $SE = .09$) did not appear to differ in their openness ratings ($MD = 0.095$, $SE = .097$, $p_{bonferroni} = 1$; $t(504) = 0.98$, $p = .327$, $d = 0.04$), whereas for females, feminine faces ($M = 5.26$, $SE = .09$) were rated as higher on openness compared to masculine faces ($M = 4.17$, $SE = .09$; $MD = -1.081$, $SE = .114$, $p_{bonferroni} = .463 t(504) = -9.48$, $p < .001$, $d = -0.42$; Figure 5.17.).

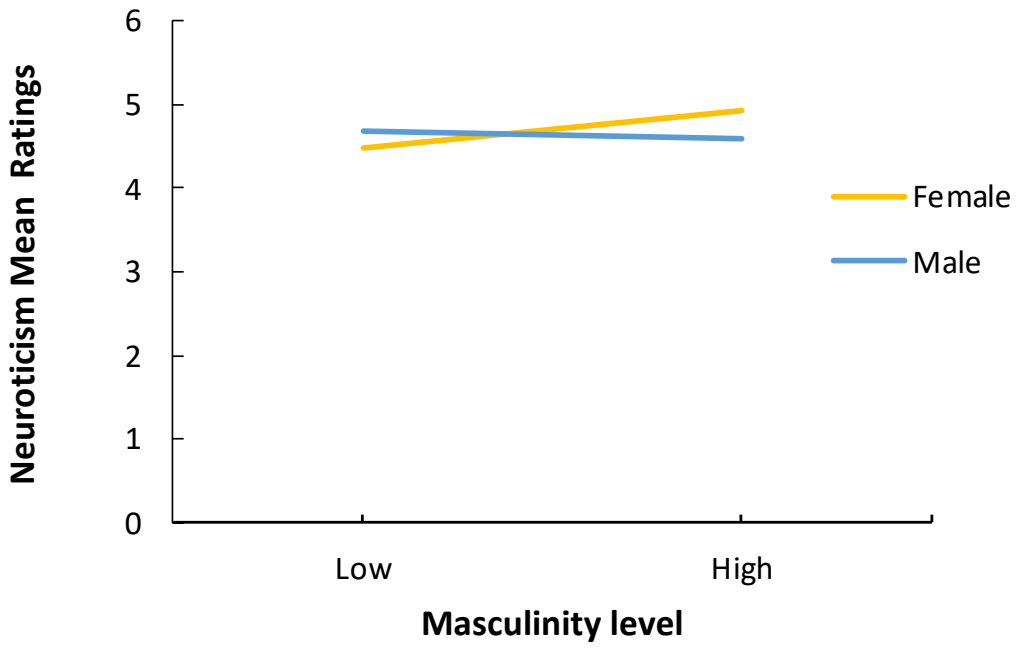


Figure 5.18. Interaction Between Masculinity Level and Gender on Neuroticism Ratings

For Neuroticism, results showed that there was a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 10.802$, $p < .01$, $\eta^2 = .02$. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for male faces, feminine faces ($M = 4.67$, $SE = .09$) and masculine faces ($M = 4.59$, $SE = .09$) did not significantly differ in neuroticism ratings ($MD = -0.083$, $SE = .116$, $p_{bonferroni} = 1$; $t(504) = -0.71$, $p = .473$, $d = -0.03$), whereas for females, masculine faces ($M = 4.93$, $SE = .1$) were rated as higher on neuroticism compared to feminine faces ($M = 4.48$, $SE = .09$; $MD = 0.452$, $SE = .12$, $p_{bonferroni} = .012$; $t(504) = 3.78$, $p < .001$, $d = 0.17$; Figure 5.18.).

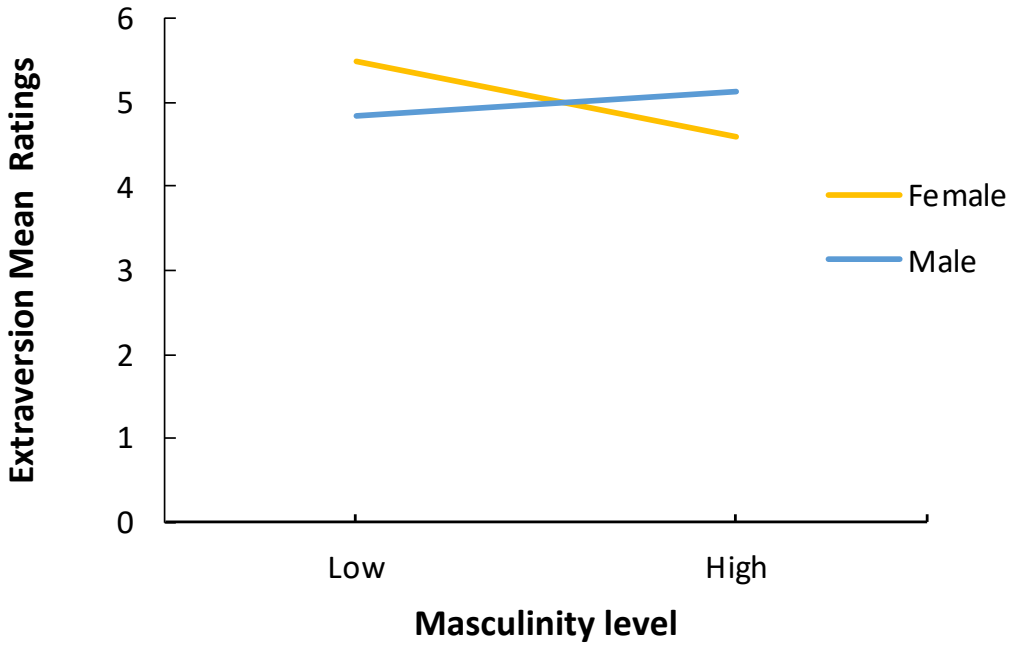


Figure 5.19. Interaction Between Masculinity Level and Gender on Extraversion Ratings

For Extraversion, results showed that there was a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 59.305$, $p < .001$, $\eta^2 = .11$. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for male faces, masculine faces ($M = 5.14$, $SE = .09$) and feminine faces ($M = 4.85$, $SE = .09$) did not significantly differ in their extraversion ratings ($MD = 0.297$, $SE = .106$, $p_{bonferroni} = .356$), although omitting the Bonferroni correction showed a significant difference in manipulation levels, $t(504) = 2.80$, $p < .01$, $d = 0.12$. On the other hand, for female faces, feminine faces ($M = 5.48$, $SE = .09$) were rated as higher on extraversion compared to masculine faces ($M = 4.58$, $SE = .1$; $MD = -0.897$, $SE = .109$, $p_{bonferroni} < .001$; $t(504) = -8.20$, $p < .001$, $d = -0.37$; Figure 5.19.).

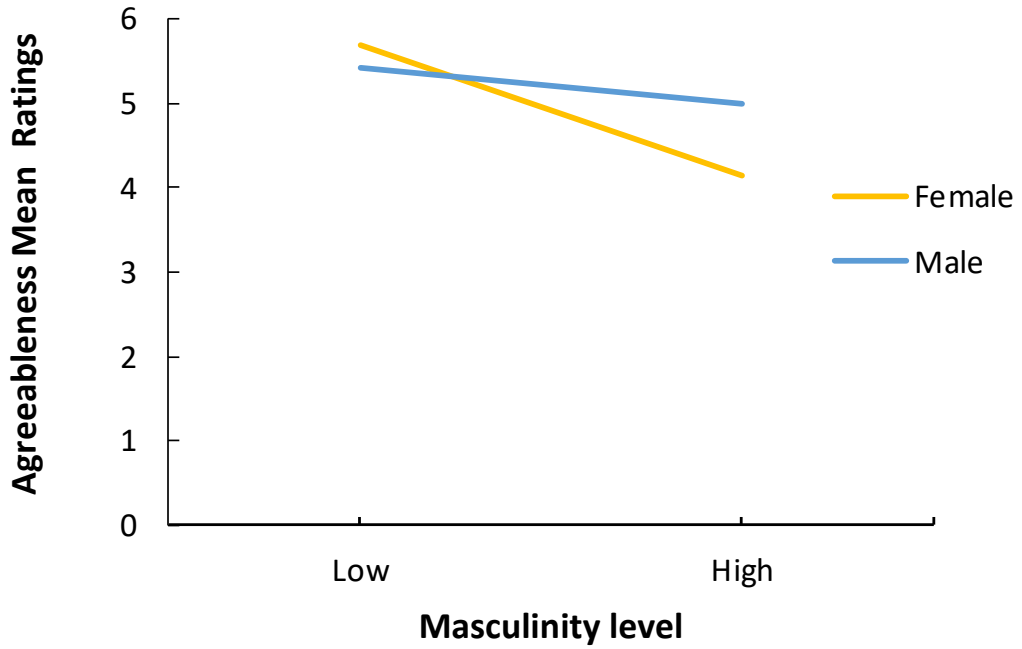


Figure 5.20. Interaction Between Masculinity Level and Gender on Agreeableness Ratings

For Agreeableness, results showed that there was a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 52.989$, $p < .001$, $\eta^2 = .10$. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for both male and female faces, feminine faces ($M_{Female} = 5.7$, $SE_{Female} = .09$; $M_{Male} = 5.42$, $SE_{Male} = .09$) were rated as higher on agreeableness compared to masculine faces ($M_{Female} = 4.14$, $SE_{Female} = .09$; $M_{Male} = 5$, $SE_{Male} = .09$). However, this difference in agreeableness rating based on masculinity level was more prominent for female faces ($MD = -1.566$, $SE = .12$, $p_{bonferroni} < .001$; $t(504) = -13.01$, $p < .001$, $d = -0.58$, compared to male faces ($MD = -0.418$, $SE = .104$, $p_{bonferroni} = .004$; $t(504) = -4.03$, $p < .001$, $d = -0.18$; Figure 5.20).

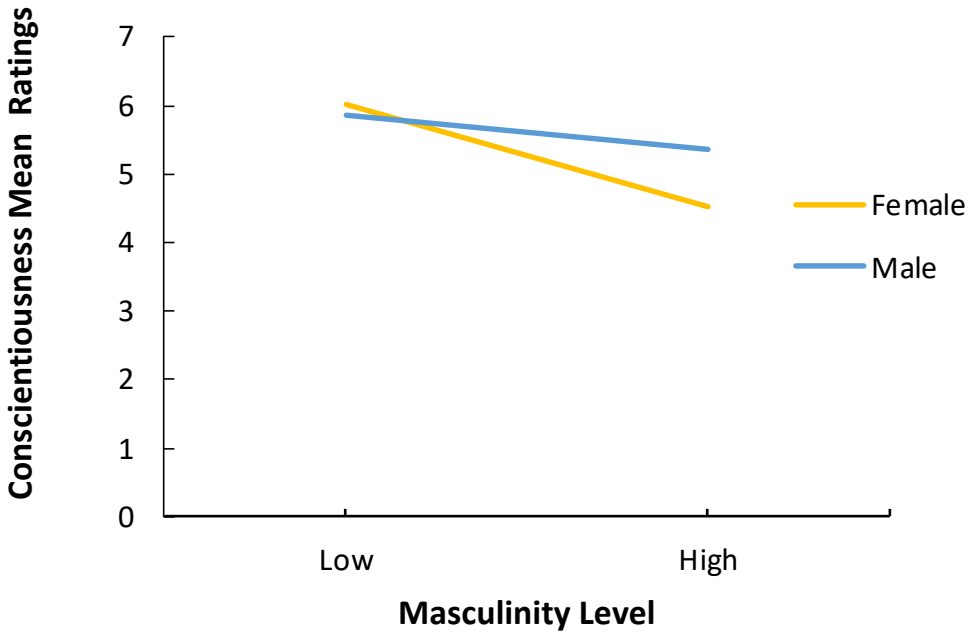


Figure 5.21. Interaction Between Masculinity Level and Gender on Conscientiousness Ratings

For Conscientiousness, results showed that there was a statistically significant interaction between Manipulation Level and Gender, $F(1, 504) = 49.904$, $p < .001$, $\eta^2 = .09$. Namely, post hoc tests using Bonferroni corrected paired samples t-test revealed that, for both male and female faces, feminine faces ($M_{Female} = 6$, $SE_{Female} = .08$; $M_{Male} = 5.86$, $SE_{Male} = .08$) were rated as higher on conscientiousness compared to masculine faces ($M_{Female} = 4.51$, $SE_{Female} = .09$; $M_{Male} = 5.37$, $SE_{Male} = .09$). However, this difference in conscientiousness rating based on masculinity level was more prominent for female faces ($MD = -1.493$, $SE = .1$, $p_{bonferroni} < .001$; $t(504) = -14.96$, $p < .001$, $d = -0.67$) compared to male faces ($MD = -0.487$, $SE = .107$, $p_{bonferroni} < .001$; $t(504) = -4.54$, $p < .001$, $d = -0.20$; Figure 5.21).

CHAPTER 6 DISCUSSION

This thesis aimed to add to the literature on face perception and impression formation by extending previous research conducted on the effects of face features on trustworthiness and dominance impressions to more specific personality trait impressions. Mainly, through this pre-registered study, it was investigated how (1) facial skin smoothness, (2) masculinity, and (3) baby facedness influence judgements about others' Big Five and Dark Triad traits. To answer these questions, faces were manipulated such that there was a high and low level for each face category (exp, high skin smoothness vs. low skin smoothness) and it was assessed whether participants' personality trait inferences differed in response to the manipulation levels.

Overall findings showed that manipulating levels of facial skin smoothness, baby facedness and masculinity significantly affected participants' personality trait impressions. Thus, suggesting that individuals rely on these specific facial features when making judgements about others' Big Five and Dark Triad traits. Upon closer examination, it is revealed that facial skin smoothness, masculinity, and baby facedness influence different traits in distinct ways. In the next section, the findings of all three face feature categories are discussed in more detail together with possible explanations for the findings, contributions to the literature, limitations of the study and suggestions for future directions.

6.1. POSSIBLE EXPLANATIONS OF FINDINGS

6.1.1. The Effect of Skin Smoothness

In relation to skin smoothness, based on the literature (see chapter 2.7.1.) it was expected that faces with blemished skin would be judged as having a higher score on more negative traits (Machiavellianism, Psychopathy and Neuroticism; H1) whereas faces with smoothed skin would be judged as having a higher score on more positive traits (Extraversion, Agreeableness, Conscientiousness, Openness) and Narcissism (H1). Results revealed that the expectations above were only met for two out of the

eight traits: Narcissism and Extraversion. Namely, faces with smoothed skin were rated as higher on both Narcissism and Extraversion compared to faces with blemished skin. Previous research has found that observer inferences of attractiveness are positively associated with inferences of Extraversion (Langlois et al., 2000) and that, additionally, there is also a small but reliable positive narcissism–attractiveness correlation in observers' ratings (Holtzman & Strube, 2010; Rauthmann & Kolar, 2012). Since skin smoothness is positively correlated with attractiveness (Fink et. al., 2001; Tsankova & Kappas, 2016; Jaeger et al., 2017), the mentioned findings regarding Extraversion and Narcissism fit well within the literature and show that previously reported results of the effect of attractiveness also extend to skin smoothness (Langlois et al., 2000; Holtzman & Strube, 2010; Rauthmann & Kolar, 2012).

On the other hand, for the remaining six trait ratings, results either revealed no notable difference in trait ratings between the two levels of skin smoothness (Neuroticism, Openness, Agreeableness, Conscientiousness) or revealed a difference that was in the opposite direction of what was expected (Psychopathy, Machiavellianism; smoothed skin faces were rated as higher on psychopathy and Machiavellianism compared to blemished skin faces). The first finding would simply suggest that individuals did not rely on skin smoothness when making judgements about others' level of Neuroticism, Openness, Agreeableness and Conscientiousness. However, the latter finding is worth discussing in more detail: Why was the opposite of what was expected found in the data for Psychopathy and Machiavellianism?

It could be possible that the smoothed skin faces were perceived as fake and inauthentic, thus more likely to be associated with the deceitful nature of the dark triad traits (Paulhus & Williams, 2002). In fact, an interesting finding in the dark triad literature is that these traits have been linked to short-term mating success (Jonason et al., 2010, Jonason et al., 2009; Paulhus & Williams, 2002), which has been previously interpreted as a by-product of these individuals' proneness to deception. Relatedly, this finding can also be linked to previous research showing that impressions of dark triad traits are positively associated with perceived attractiveness (Carter et al., 2014; Holtzman & Strube, 2012). Researchers have suggested that dark personalities construct certain favourable appearances that can act as social lures (Holtzman & Strube, 2012), but also that individuals, in particularly woman, consciously rate targets described as having high dark triad traits as having a more attractive personality

(Carter et al., 2014). These two claims contradict one another, however, they both support the underlying message that, although Psychopathy, Machiavellianism and Narcissism consist of negative traits, they are positively responded to in terms of attractiveness. With these points in mind while also remembering the positive association between skin smoothness an attractiveness (Fink et. al., 2001; Tsankova & Kappas, 2016; Jaeger et al., 2017), my finding of smoothed skin faces being rated as higher on Psychopathy and Machiavellianism compared to blemished faces is perhaps not all too unexpected. To the best of my knowledge, so far, no research has assessed the influence of skin smoothness on inferences about the dark triads, therefore, future work should analyse these relations in order to produce more comparable results and to be able to make stronger inferences.

Further, based on previous reports of an interaction between gender and attractiveness on perceptions of specific personality traits (Extraversion, Conscientiousness and Openness; Tartaglia & Rollero, 2015; Ćurković & Franc, 2010; Langlois et al., 2000), it was also explored whether such a gender interaction extends to skin smoothness in relation to impressions of the eight personality traits (EQ1). Results revealed a significant interaction between gender and skin smoothness for Psychopathy, Machiavellianism, Narcissism, Extraversion and Agreeableness impressions. However, no interaction effect of gender and skin smoothness were found for impressions of Openness, Neuroticism and Conscientiousness.

Namely, the results revealed that gender differences in attractiveness-based extraversion impressions (Tartaglia and Rollero, 2015) also conceptually replicate for perceived skin smoothness: Smoothed skin faces were considered to be more extraverted than blemished faces, yet this effect of skin smoothness was only significantly present for male faces. However, previously found gender differences in attractiveness-based Conscientiousness and Openness impressions (Tartaglia & Rollero, 2015) did not replicate for perceived skin smoothness. For conscientiousness, previously reported gender difference in Conscientiousness ratings (irrespective of skin smoothness) was also present in the current study (Tartaglia & Rollero, 2015), such that female faces were rated as higher on Conscientiousness compared to male faces ($F(1, 504) = 133.59$, $p < .001$). However, no such gender differences were found on the effect of skin smoothness on Conscientiousness ratings. This is likely due to the study's finding of no significant influence of skin smoothness on Conscientiousness ratings. Further, contrary to Tartaglia and Rollero (2015), there was a main effect of

gender on openness ratings ($F(1, 504) = 15.10$, $p < .001$) such that female faces were rated as higher on Openness compared to male. However, again contrary to Tartaglia and Rollero (2015), no gender differences were found on the effect of skin smoothness, again, likely due to the lack of influence that skin smoothness had on Openness ratings. Another, yet only partly, contradicting finding to the literature was in relation to Agreeableness (Ćurković & Franc, 2010; Langlois et al., 2000; Tartaglia & Rollero, 2015): While previous work found no interaction effect of gender and attractiveness, the current study found a significant interaction between gender and skin smoothness on Agreeableness impressions. However, when post hoc analyses were conducted, it was revealed that both male and female faces did not actually significantly differ in their agreeableness ratings in response to skin smoothness levels, which was in line with previous findings in the literature (Ćurković & Franc, 2010; Langlois et al., 2000; Tartaglia & Rollero, 2015). For Neuroticism, findings were conceptually in line with Tartaglia and Rollero (2015), such that there was no interaction effect of gender and skin smoothness on Neuroticism ratings.

Importantly, to the best of my knowledge, this study is the first to report that Dark Triad trait impressions based on facial skin smoothness differ for male and female faces. In relation to Psychopathy: Female faces did not appear to be influenced by skin smoothness, however, for male faces, smoothed faces were rated as higher on psychopathy compared to blemished faces. The same was the case for Machiavellianism: In female faces, smoothed and blemished faces did not appear to notably differ in their Machiavellianism ratings, whereas in male faces, smoothed faces were rated as higher on Machiavellianism compared to blemished faces Finally, for Narcissism: Female smoothed and blemished faces did not appear to notably differ in their Narcissism ratings, whereas for males, smoothed faces were rated as higher on Narcissism compared to blemished faces. Since there is no existing comparable work in the literature and since these findings were based on explorative questions, future work should try to see if these gender differences in the effect of skin smoothness do in fact replicate.

6.1.2. The Effect of Baby Facedness

In relation to baby facedness, based on the literature (see chapter 2.7.2.) it was expected that babyish faces (i.e., high level baby facedness) would be judged as having a higher score on agreeableness whereas mature faces (i.e., low level baby facedness)

would be judged as having a higher score on Conscientiousness, Psychopathy, Machiavellianism and Narcissism. Results revealed that my expectations were met for four out of the five traits: Psychopathy, Narcissism, Machiavellianism and Agreeableness. Mature faces (i.e., low level baby facedness) were rated as higher on Psychopathy, Machiavellianism and Narcissism compared to babyish faces (i.e., high level skin smoothness). Since the Dark Triad traits have previously been found to be related to perceptions of untrustworthiness (Rogers et al., 2018; Gordon & Platek, 2009), this finding supports previous research showing that baby facedness is overall associated with being warm, honest, trustworthy but incompetent (Montepare & Zebrowitz, 1998; Zebrowitz & Montepare, 1992). Thus, the finding of attributing lower dark triad trait ratings to babyish faces is fitting with the literature. Further, it was found that babyish faces were rated as higher on Agreeableness compared to mature faces, which compliments previous work showing that babyish features are associated with being naïve and submissive (Montepare & Zebrowitz, 1998; Zebrowitz & Montepare, 1992).

For the remaining trait, Conscientiousness, results revealed no notable difference in trait ratings between the two levels of baby facedness. The lack of effect of baby facedness on Conscientiousness ratings is quite puzzling because previous research shows that individuals with babyish features have been found to be avoided for assigning mentally challenging tasks and leadership positions (Montepare & Zebrowitz, 1998). On the other hand, Conscientiousness is one of the most desired personality traits in the workplace (Wilmot & Ones, 2019), associated with emerging as a leader (Wilmot & Ones, 2019) and characterized as being hardworking, organized, and careful (Robert et al., 2009). Thus, one would expect that mature faces would be assigned higher conscientiousness ratings compared to babyish faces. Perhaps, the lack of a significant difference in babyish and mature faces for conscientiousness impressions could be attributed to the manipulation not being strong enough, although that possibility would suggest that Psychopathy, Machiavellianism, Narcissism and Agreeableness ratings also should not differ in response to the employed high and low baby facedness faces. Therefore, the results of the current study would suggest that individuals do not rely on facial baby facedness when making assumptions about an individual's conscientiousness level, though future research should continue to assess whether in fact this is the case.

Further, since no study has previously assessed the influence of babyish features perceptions of Extraversion, Openness and Neuroticism, this thesis also sought to explore whether babyish and mature faces differ in their attributed Extraversion, Openness and Neuroticism ratings (EQ2). Results showed that babyish faces were rated as significantly higher on openness and extraversion compared to mature faces. However, for neuroticism, babyish and mature faces did not appear to significantly differ in their attributed ratings.

Since Extraversion has been previously reported as the most consistent correlate of leadership (Judge et al., 2002) and babyish features are generally avoided for leadership positions (Montepare & Zebrowitz, 1998), the finding that babyish faces were associated with higher extraversion appears to contradict the literature. Perhaps, the babyish faces were perceived as more approachable and thus also more extraverted (Sutherland et al., 2015). The finding that babyish faces were also rated as more agreeable supports this possibility, however more research is needed to make stronger inferences.

For Neuroticism and Openness, to the best of my knowledge, there does not exist any previous work to which the findings of the current study can be compared to. The results suggest that people are indeed affected by baby facedness when making inferences about others' openness and that they do not rely on this facial feature when forming impressions about neuroticism, yet more research is needed to better discuss and interpret these findings.

6.1.3. The Effect of Masculinity

In relation to masculinity, based on the literature (see chapter 2.7.3.) it was expected that masculine faces (i.e., high level masculinity) would be judged as having a higher score on the Dark Triad traits (Psychopathy, Machiavellianism and Narcissism) whereas feminine faces (i.e., low level masculinity) would be judged as having a higher score on agreeableness and openness. Results revealed that my expectations were met for all five traits: Psychopathy, Machiavellianism, Narcissism, Openness and Agreeableness. For the dark triad traits, this finding supports previous work showing that masculine faces are generally perceived to be less trustworthy when compared with feminine faces (Perett et. al., 1998; Walker & Wänker, 2017). Since the Dark Triad traits have also been found to be associated with perceptions of untrustworthiness (Rogers et al., 2018; Gordon & Platek, 2009), it is fitting that more

masculine faces were attributed higher ratings of Psychopathy, Machiavellianism and Narcissism in this study. Further, Walker and Vetter's (2016) revealed that perceptions of agreeableness and openness to experience strongly overlap with trustworthiness and have a strong negative correlation with dominance. Thus, since masculine faces are associated with untrustworthiness and dominance (Perett et. al., 1998; Walker & Wänker, 2017; Walker & Vetter, 2016), the finding that feminine faces were attributed a higher score on agreeableness and openness compared to masculine faces, again, supports previous results found in the literature.

So far, no study has looked at the influence of masculine face features on perceptions of Neuroticism, Extraversion and Conscientiousness. Thus, this thesis aimed to also explore whether masculine and feminine faces differ in their attributed Neuroticism, Extraversion and Conscientiousness ratings (EQ3). Results revealed that, for neuroticism, faces with high and low levels of facial masculinity did not significantly differ in their trait ratings. On the other hand, for extraversion and conscientiousness, feminine faces were rated as significantly higher on these traits compared to masculine faces.

Finally, based on previous research showing that the effect of facial masculinity on trait impressions is more pronounced for atypical (i.e., masculine-looking women and feminine-looking men) faces (Walker & Wänker, 2017), the possible interaction between gender and facial masculinity was explored (EQ4). Results showed that there was a significant gender and facial masculinity interaction for all personality trait impressions (Psychopathy, Machiavellianism, Narcissism, Openness, Neuroticism, Extraversion, Agreeableness and Conscientiousness).

For both male and female faces, masculine faces were rated as higher on Psychopathy, Machiavellianism and Narcissism compared to feminine faces, however this effect of facial masculinity was more prominent for female faces. This suggests that the effect of facial masculinity (i.e., masculine faces being rated as higher in the dark triad traits) is stronger for gender-atypical faces of women than for gender-atypical faces of men.

In regard to Neuroticism, typical and a-typical male faces did not differ in their trait ratings, whereas for female faces, masculine (a-typical) female faces were rated as higher on Neuroticism compared to feminine (typical) female faces. On the other hand, regarding Extraversion, while gender typical and a-typical male faces did not differ in their trait ratings, gender-typical female faces were rated as higher in

extraversion compared to gender-atypical female faces. For Openness, it was found that the effect of facial masculinity differed for gender-atypical women and men. Namely, for openness, gender-typical and gender-atypical male faces did not appear to notably differ in their attributed ratings, whereas gender-typical female faces were rated as higher on openness compared to gender-atypical female faces. Finally, in terms of agreeableness and conscientiousness, for both males and females, feminine faces were rated as higher on agreeableness and conscientiousness compared to masculine faces. However, this difference in agreeableness and conscientiousness ratings based on masculinity level was more prominent for female faces, suggesting that the effect of facial masculinity (i.e., feminine faces being rated as higher in agreeableness and conscientiousness) is stronger for gender-atypical faces of women than for gender-atypical faces of men. Since the current study is the first to report interactions between gender and facial masculinity on inferences about the dark triad and big five personality traits, more research is needed to provide a better interpretation of these findings.

6.1.4. Evolutionary Based Explanations

How come there is consensus on certain facial features being associated with specific (Big five and Dark Triad) personality trait impressions? More cross-cultural work is required to really be able to paint a picture of how much we actually agree on certain facial cues signalling certain characteristic traits. However, the findings of the current study support the idea that there are indeed implicit stereotypes about what people with specific personality traits look like. One way to explain such consensus is via evolutionary accounts: Our tendency to judge individuals based on their face, especially in regard to the fundamental traits warmth and dominance, provide adaptive functions for our social interactions. In particular, the overgeneralization hypothesis, which mainly developed from research regarding baby facedness (Zebrovitz et al., 2003; Zebrowitz & Montepare, 2008), posits that our impressions of people result from overgeneralizing perceived facial cues as signalling fundamental characteristics of that person, such as whether to approach/avoid that person (based on cues of valence) or how strong/weak that person is (based on cues of dominance; Oosterhof. &. Todorov, 2008). Relying on such facial cues was presumably quite beneficial in the past, we cooperated with those we saw as warm and avoided those we saw as cold. Now as the world as well as the humans in it have gotten more and more sophisticated, it is likely

that we are extending this adaptive and primal ability to more specific personality trait impressions, since it aids us in navigating through our complex social environment (such as deciding who to hire or who to sentence as guilty). Thus, the findings of the current study, communally associating high and low levels of skin smoothness, baby facedness and masculinity with differing levels of the big five and dark triad traits, can be seen as a manifestation of a broader mechanism (facial overgeneralization) that is rooted in our evolutionary adaptation.

Another relevant point worth noting is that, overall, the strength as well as range of associations between the tested face features and trait ratings were stronger and narrower for the dark triad traits (d= 0.21-0.53) compared to the Big Five Traits (d= 0.01-0.59). In other words, the influence of the face features appeared to be more consistent and effective for impressions about the dark triad traits in comparison to impressions about the big five traits, where the effects were rather more scattered along different face features and personality dimensions. From an evolutionary point of view, this pattern can again be viewed as a reflection of certain underlying adaptive abilities. Namely, according to the Error Management Theory (Haselton & Buss, 2000), when faced with uncertainty (such as trying to judge a person's character from their face), our cognition functions in a way that aims to minimize the evolutionary fitness cost of any potential errors in judgements (DeLecce, 2018). Since the potential costs of failing to recognize that a person has dark triad traits would be much more costly (for example, individuals high on dark triad traits exhibit exploitative mating strategy; Jonason et al., 2012) than failing to correctly judge a person's big five traits, it seems plausible that our face-based judgements regarding others' dark triad traits emerge in a stronger and clearer manner (compared to big five traits) due to its evolutionary adaptive benefits. Relatedly, the Self-Domestication Hypothesis (Wrangham, 2019) proposes that, similar to how we have domesticated wild animals (for example, wolves into dogs), in human evolution we have gone through a process of selection where we have selected against aggression (Sánchez-Villagra & van Schaik, 2019) and favoured pro-sociality (Hare, 2017). Recent work shows support for this hypothesis in the modern human face, suggesting that, in the past, certain faces (that looked more friendly and less aggressive than others) were likely favoured by humans which lead to the current appearance of human faces (Zanella et al.,2019). Thus, in relation to this thesis, it could be argued that the finding of the dark triad traits being more clearly and strongly associated with certain facial feature compared to the

big five traits may stem from the possibility that we are programmed to detect (and select against) faces that are more aggressive and costly for us.

6.2. CONTRIBUTIONS OF THE STUDY

We know that individuals rely on people's facial appearance when making judgements about their character. However, you cannot actually determine the character of a person just by looking at their face. The research so far suggests that, at most, we are only moderately accurate in judging others' personalities from their faces (Walker & Vetter, 2016), leaving quite a lot of room for errors. Alarmingly though, not only is there consensus in regard to which faces signal which traits (Walker & Vetter, 2016), people actually believe that they can accurately judge a person's character based on their face (Jaeger et al., 2020) and act upon these incorrect judgements (Olivola & Todorov, 2010; Todorov et al., 2015). Regarding the big five and dark triad traits, previous work has mainly focused on individuals' accuracy in face-based impressions, however, perhaps a more critical question to ask is: What are individuals' beliefs and stereotypes about how individuals who vary in their big five and dark triad traits look like? Because whether we are accurate or not, it is such beliefs and stereotypes that guide our decisions, behaviour and influence social outcomes (Olivola et al., 2014b; Todorov et al., 2015). The current study is one of the first attempts to address this knowledge gap. Namely, by directly manipulating certain facial features (skin smoothness, masculinity, baby facedness) and avoiding the use of composite images (as previously preferred; Alper et al., 2021; Holtzman, 2011; Penton-Voak et al., 2006; Little & Perett, 2007), the current study provides direct insight into how the assessed facial features are associated with perceptions of the big five and dark triad traits.

Secondly, the research conducted on this topic have been based on WEIRD samples (Heinrich et al., 2010; Walker & Vetter, 2016; Walker et al., 2018). As of yet, no study has assessed how facial features affect Turkish individuals' impressions of others' big five and dark triad traits. By utilizing a Turkish sample, considered to be a non-WEIRD population, this study reveals to what extent individuals from the Turkish culture rely of certain face features (skin smoothness, masculinity, baby facedness) when making inferences about specific personality traits. Diversifying the samples in which face-based impressions are studied upon will provide more insight into the

possible evolutionary background of our superficial personality trait inferences based on faces.

6.3. LIMITATIONS

Together with its contributions the current study also has its limitations. Firstly, one possible limitation of the study is that participants rated all eight personality traits simultaneously. Initially, it was planned for each face to be rated on a single personality trait and for each personality trait to be accompanied by a different face identity, resulting in a total of 128 faces to be rated. However, upon conducting the pilot study, it was quickly realised that this resulted in a survey that was too long and exhaustive to reach a sufficient sample size and to preserve participants' attention. Thus, it was decided that instructing participants to rate each face on all eight personality traits was a more appropriate method of choice. Via this method, it is likely that the initial personality trait rated for each face affected all following trait ratings for the same face. Ideally, the use of randomized order for all trait ratings should cancel out any possibility of order effects. Further, previous researchers have provided support for a Big-Two model of personality (DeYoung, 2006; DeYoung et al., 2002; Digman, 1997), mainly suggesting that the Big five dimensions are not independent from another and research regarding the Dark Triad traits has found that Psychopathy, Machiavellianism and Narcissism are consistently and positively correlated with one another (Jonason et al., 2020). These two points could suggest that the utilized method for assessing face-based impressions may not actually interfere with the nature of how personality exhibits itself out in the real world: Personality dimensions are not independent, rather they exist relative to one another (Digman, 1997). Nevertheless, future studies should look into whether any different procedures of testing face-based trait impressions, such as having participants rate each separate face identity on a single personality trait, result in stronger associations or different outcomes.

A second limitation is that, in the pilot study assessing whether the faces manipulated to be high and low in their respective face features were indeed perceived that way, each face pair was presented side by side to the participants. In other words, participants saw the high level and low level face simultaneously when giving their judgements of skin smoothness, baby facedness and masculinity for each face. It is possible that presenting the faces (of each pair) side by side motivated participants to specifically look for differences in the two faces and, thus, somewhat inflated the

perceived difference between the high and low level faces. Future work could investigate whether presenting the faces sequentially drastically changes the results of such a face manipulation check in comparison to when they are presented side by side.

Another limitation concerns the sample of the current study. Namely, most of the sample consisted of female participants. Especially in regard to the trait inferences of masculine and feminine faces, this lack of a normal distribution in gender may have created a bias. Previous work shows that male and female observers do indeed differ in face perception, for example, both genders ascribe their positive stereotypical attributes to same gender faces such as men seeing male faces as stronger than women do (Bracci et al., 2021). Thus, it is possible that the findings of the current study only represent the consensus regarding face-based impressions among a specific group of individuals, that is, women. Future research should aim to see whether these results generalize across both genders and, further, whether there are any gender differences regarding face-based impressions of the Big Five and Dark Triad traits.

A further limitation of this study is that no attention checks were included that could discern whether participants randomly assigned trait ratings to the faces or whether they indeed relied on the faces while making their judgements. In other words, the inclusion of an attention check could have provided a better understanding of the current studies data quality. To try to compensate for this limitation, in addition to the main analyses, two additional sets of paired samples t-test looking at the effect of the face features on trait ratings were conducted (see Appendix B) in which only participants who finished the study in over 7 minutes (N =478) and 10 minutes (N =393) were included. Overall results of this comparison showed that excluding individuals who quickly responded to the items of the questionnaire did not change the results of the main findings in any meaningful ways. For details see Appendix B.

Further, did individuals truly rely on the facial features that were manipulated? This question addresses an empirical limitation of the current study. As discussed by Jaeger and Jones (2021), examining only one feature or a few features in isolation makes it difficult to discern whether individual truly relied on the facial feature of interest, or on other correlated features. By only manipulating the face feature in question and keeping all other aspects of the face constant, one would expect that any variance in response to the face would be driven by the face feature manipulation. This thesis provides one of the many ways forward in assessing Big Five and Dark Triad trait impressions in relation to face perception. However, where possible, having

experimental control over all facial features, which is usually obtained via machine learning methods (Jaeger & Jones, 2021), would provide stronger evidence for which facial features are related to which big five and dark triad traits.

The current study did not include any measures of confidence. That is to say, although this study provides insight into individuals' beliefs and stereotypes about what, for example, an extravert or psychopath looks like, no conclusions can be drawn on how confident individuals are in their face-based judgements of the big five and dark triad traits. The literature suggests that individuals believe their judgements to be true (Jaeger et al., 2022; Willis & Todorov, 2006) and that such confidence increases with exposure time even though the first impressions do not change much (Willis & Todorov, 2006). Future studies should incorporate measures of confidence to get a better state of the art and to see whether confidence varies based on the trait in question. For instance, guessing a person's level of extraversion based on their face may be perceived as easier than guessing their level of psychopathy.

Finally, due to the aim of looking for consensus rather than any individual differences in judgements, personality ratings were aggregated across participants. However, this method unfortunately masks the role of any stable individual differences in face-based impression. Future work should try to establish to what extent individuals differ in their reliance on facial feature when inferring the big five and dark triad traits as well as to what extent they differ on which facial features signal which traits.

6.4. FUTURE DIRECTIONS

Does providing a time constraint change individuals' reliance on these facial features in question? Previous work shows that increasing individuals' exposure time to a face from 500 to 1,000 ms does not drastically change their judgements of basic traits such trustworthiness and competence (Willis & Todorov, 2006). No research has been conducted on whether the same is true for impressions of more specific traits such as the Big Five and Dark Triad traits. Further, a future direction could be to look at whether putting participants into different face exposure conditions (ranging from 50ms to 1000ms) changes the effect of facial skin smoothness, masculinity, and baby facedness on impressions of the Big Five and Dark Triad traits. No significant changes in effect would signal that the face feature in question is quite primal in forming such impressions.

To what extent does the personality of the observer influence their face-based judgements of others' Big Five and Dark Triad traits? Previous research has found that some personality traits of the observer, such as neuroticism, extraversion, openness, conscientiousness influence how they perceive a face to be (Bracci et al., 2021). Further, according to the phenomenon of assumed similarity (Cronbach, 1955; Kenny & West, 2010), people show a tendency to perceive others, even random strangers, as somewhat similar to them. In other words, to some extent, they tend to rely on their own personality characteristics when judging the personality of others (Thielmann & Hilbig, 2022) Such a bias is not grounded in any actual level of similarity but simply just an illusion (Human & Biesanz, 2011; Thielmann & Hilbig, 2022). Future work should include measures of the observers' personality traits to assess to what extent the observers'' personality and the phenomenon of assumed similarity moderate the effect of facial features (such as skin smoothness, masculinity and baby facedness) on impressions of Big Five and Dark Triad traits.

As a final note, to keep the experiment manageable, the current study analysed only three (skin smoothness, masculinity and baby facedness) out of a variety of different face features that can be studied (Jeager & Jones, 2021). An important future direction would be to assess the influence of other face features as well, such as facial width to height ratio, race, emotion resemblance, familiarity, etc. (Jeager & Jones, 2021) to gain a better understanding of our stereotypes and beliefs regarding how the Big Five and Dark Triad traits look like.

All materials and data can be provided by the author upon request (busraelifyelbuz@gmail.com) for any research purposes.

REFERENCES

Alper, S., Bayrak, F., & Yilmaz, O. (2021). All the Dark Triad and some of the Big Five traits are visible in the face. *Personality and Individual Differences*, *168*, 110350.

Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play!. *Science*, *323*(5918), 1183-1183.

Balaban, M. T. (1995). Affective influences on startle in five-month-old infants: Reactions to facial expressions of emotion. *Child development*, *66*(1), 28-36.

Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences*, *104*(46), 17948-17953.

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, *6*(2), 269.

Behrmann, M., Richler, J. J., Avidan, G., & Kimchi, R. (2015). Holistic face perception. *Oxford handbook of perceptual organization*, 758-774.

Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces: the impact of age-related craniofacial changes on social perception. *Psychological bulletin*, *100*(1), 3.

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

Borkenau, P., Brecke, S., Möttig, C., & Paelecke, M. (2009). Extraversion is accurately perceived after a 50-ms exposure to a face. *Journal of research in personality*, *43*(4), 703-706.

Bracci, M., Guidi, S., Marchigiani, E., Masini, M., Palmitesta, P., & Parlangeli, O. (2021). Perception of Faces and Elaboration of Gender and Victim/Aggressor Stereotypes: The Influence of Internet Use and of the Perceiver's Personality. *Frontiers in psychology*, *12*, 2014.

Buckingham, G., DeBruine, L. M., Little, A. C., Welling, L. L., Conway, C. A., Tiddeman, B. P., & Jones, B. C. (2006). Visual adaptation to masculine and feminine faces influences generalized preferences and perceptions of trustworthiness. *Evolution and Human Behavior*, *27*(5), 381-389.

Carter, G. L., Campbell, A. C., & Muncer, S. (2014). The dark triad personality: Attractiveness to women. *Personality and Individual Differences*, *56*, 57-61.

Castelli, L., Carraro, L., Ghitti, C., & Pastore, M. (2009). The effects of perceived competence and sociability on electoral outcomes. *Journal of Experimental Social Psychology*, *45*(5), 1152-1155.

Chang, L. J., Doll, B. B., van't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive psychology*, *61*(2), 87-105.

Chen, F. F., Jing, Y., & Lee, J. M. (2014). The looks of a leader: Competent and trustworthy, but not dominant. *Journal of Experimental Social Psychology*, *51*, 27-33.

Chiao, J. Y., Bowman, N. E., & Gill, H. (2008). The political gender gap: Gender bias in facial inferences that predict voting behavior. *PloS one*, *3*(10), e3666.

Citrin, L. B., & Roberts, T. A. (2004). A feminist psychological perspective on gendered affect. *The social life of emotions*, 203-223.

Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from faces: A developmental study. *Psychological science*, *25*(5), 1132-1139.

Conway, C. A., Jones, B. C., DeBruine, L. M., & Little, A. C. (2010). Sexual dimorphism of male face shape, partnership status and the temporal context of relationship sought modulate women's preferences for direct gaze. *British Journal of Psychology*, *101*(1), 109-121.

Cronbach, L. J. (1955). Processes affecting scores on" understanding of others" and" assumed similarity.". *Psychological bulletin*, *52*(3), 177.

Ćurković, K., & Franc, R. (2010). Physical attractiveness stereotype'Beautiful is Good'in the context of the Big Five personality theory. *Psihologijske teme*, *19*(1), 123-144.

DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., ... & Tiddeman, B. P. (2006). Correlated preferences for facial

masculinity and ideal or actual partner's masculinity. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1592), 1355-1360.

DeLecce, T. (2018). *Error Management Theory.* In: Zeigler-Hill, V., Shackelford, T. (eds) Encyclopedia of Personality and Individual Differences. Springer, Cham.

DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of personality and social psychology*, 91(6), 1138.

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health?. *Personality and Individual Differences*, 33, 533–552.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of personality and social psychology*, 73(6), 1246.

Fink, B., Grammer, K., & Thornhill, R. (2001). Human (Homo sapiens) facial attractiveness in relation to skin texture and color. *Journal of Comparative Psychology*, *115*(1), 92.

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

Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current directions in psychological science*, *27*(2), 67-73.

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

Gibson, J. J. (1979). The theory of affordances. The ecological approach to visual perception. In *The People, Place and, Space Reader* (pp. 56-60). Routledge New York and London.

Gordon, D. S., & Platek, S. M. (2009). Trustworthy? The brain knows: Implicit neural responses to faces that vary in dark triad personality characteristics and trustworthiness. *Journal of Social, Evolutionary, and Cultural Psychology*, *3*(3), 182.

Graham, J. R., Harvey, C. R., & Puri, M. (2017). A corporate beauty contest. *Management Science*, *63*(9), 3044-3056.

Hare, B. (2017). Survival of the friendliest: Homo sapiens evolved via selection for prosociality. *Annual review of psychology*, 68(1), 155-186.

Haselton, M. G., & Buss, D. M. (2000). Error management theory: a new perspective on biases in cross-sex mind reading. *Journal of personality and social psychology*, 78(1), 81.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and brain sciences*, *33*(2-3), 61-83.

Holtzman, N. S. (2011). Facing a psychopath: Detecting the dark triad from emotionally-neutral faces, using prototypes from the Personality Faceaurus. *Journal of Research in Personality*, *45*(6), 648-654.

Holtzman, N. S., & Strube, M. J. (2010). Narcissism and attractiveness. *Journal of Research in Personality*, *44*(1), 133-136.

Human, L. J., & Biesanz, J. C. (2011). Through the looking glass clearly: accuracy and assumed similarity in well-adjusted individuals' first impressions. *Journal of Personality and Social Psychology*, *100*(2), 349.

Jaeger, B., & Jones, A. L. (2020). Which facial features are central in impression formation?. *Social Psychological and Personality Science*, 19485506211034979.

Jaeger, B., Evans, A. M., Stel, M., & van Beest, I. (2022). Understanding the role of faces in person perception: Increased reliance on facial appearance when judging sociability. *Journal of Experimental Social Psychology*, *100*, 104288.

Jaeger, B., Wagemans, F. M., Evans, A. M., & van Beest, I. (2017). Effects of facial skin smoothness and blemishes on trait impressions. *Perception*, *47*(6), 608-625.

Jonason, P. K., Koenig, B. L., & Tost, J. (2010). Living a fast life. *Human Nature*, *21*(4), 428-442.

Jonason, P. K., LI, N. P., Webster, G. D., & Schmitt, D. P. (2009). The Dark Triad: traits that facilitate short-term mating in men. *European journal of personality*, *23*(1), 5-18.

Jonason, P. K., Luevano, V. X., & Adams, H. M. (2012). How the Dark Triad traits predict relationship choices. *Personality and Individual Differences*, 53(3), 180-184.

Jonason, P. K., Żemojtel-Piotrowska, M., Piotrowski, J., Sedikides, C., Campbell, W. K., Gebauer, J. E., ... & Yahiiaev, I. (2020). Country-level correlates of the dark triad traits in 49 countries. *Journal of personality*, 88(6), 1252-1267.

Jones, B. C., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., Arinze, N. C., ... & Sirota, M. (2021). To which world regions does the valence–dominance model of social perception apply?. *Nature human behaviour*, *5*(1), 159-169.

Jones, B. C., DeBruine, L. M., Main, J. C., Little, A. C., Welling, L. L., Feinberg, D. R., & Tiddeman, B. P. (2010). Facial cues of dominance modulate the short-term gaze-cuing effect in human observers. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1681), 617-624.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: a qualitative and quantitative review. *Journal of applied psychology*, *87*(4), 765.

Keating, C. F., & Bai, D. L. (1986). Children's attributions of social dominance from facial cues. *Child Development*, 1269-1276.

Kenny, D. A., & West, T. V. (2010). Similarity and agreement in self-and other perception: A meta-analysis. *Personality and Social Psychology Review*, *14*(2), 196-213.

Köhler, W. (1967). Gestalt psychology. *Psychologische Forschung*, *31*(1), XVIII-XXX.

Kramer, R. S., & Ward, R. (2010). Internal facial features are signals of personality and health. *The Quarterly Journal of Experimental Psychology*, *63*(11), 2273-2287.

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

Langlois, J. H., Roggman, L. A., & Rieser-Danner, L. A. (1990). Infants' differential social responses to attractive and unattractive faces. *Developmental psychology*, *26*(1), 153.

Laustsen, L. (2014). Decomposing the relationship between candidates' facial appearance and electoral success. *Political Behavior*, *36*(4), 777-791.

Lenz, G. S., & Lawson, C. (2011). Looking the part: Television leads less informed citizens to vote based on candidates' appearance. *American Journal of Political Science*, *55*(3), 574-589.

Little, A. C., & Perrett, D. I. (2007). Using composite images to assess accuracy in personality attribution to faces. *British Journal of Psychology*, *98*(1), 111-126.

Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007). Facial appearance affects voting decisions. *Evolution and Human Behavior*, *28*(1), 18-27.

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

Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach-and avoidance-related behaviors. *Emotion*, *5*(1), 119.

Martin, D. S. (1978). Person perception and real-life electoral behaviour. *Australian Journal of Psychology*, *30*(3), 255-262.

Mattes, K., Spezio, M., Kim, H., Todorov, A., Adolphs, R., & Alvarez, R. M. (2010). Predicting election outcomes from positive and negative trait assessments of candidate images. *Political Psychology*, *31*(1), 41-58.

Meier, B. P., Robinson, M. D., Carter, M. S., & Hinsz, V. B. (2010). Are sociable people more beautiful? A zero-acquaintance analysis of agreeableness, extraversion, and attractiveness. *Journal of Research in Personality*, *44*(2), 293-296.

Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of cardiac anaesthesia*, *22*(1), 67.

Montepare, J. M., & Zebrowitz-McArthur, L. (1989). Children's perceptions of babyfaced adults. *Perceptual and Motor Skills*, *69*(2), 467-472.

Montepare, J. M., & Zebrowitz, L. A. (1998). Person perception comes of age: The salience and significance of age in social judgments. In *Advances in experimental social psychology* (Vol. 30, pp. 93-161). Academic Press.

Montepare, J. M., & Zebrowitz, L. A. (1998). Person perception comes of age: The salience and significance of age in social judgments. In *Advances in experimental social psychology* (Vol. 30, pp. 93-161). Academic Press.

Morrison, D., Wang, H., Hahn, A. C., Jones, B. C., & DeBruine, L. M. (2017). Predicting the reward value of faces and bodies from social perception. *PloS one*, *12*(9), e0185093.

Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of nonverbal behavior*, *34*(2), 83-110.

Olivola, C. Y., Eubanks, D. L., & Lovelace, J. B. (2014a). The many (distinctive) faces of leadership: Inferring leadership domain from facial appearance. *The Leadership Quarterly*, *25*(5), 817-834.

Olivola, C. Y., Funk, F., & Todorov, A. (2014b). Social attributions from faces bias human choices. *Trends in Cognitive Sciences*, *18*(11), 566-570.

Olivola, C. Y., Sussman, A. B., Tsetsos, K., Kang, O. E., & Todorov, A. (2012). Republicans prefer Republican-looking leaders: Political facial stereotypes predict candidate electoral success among right-leaning voters. *Social Psychological and Personality Science*, *3*(5), 605-613.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087-11092.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of research in personality*, *36*(6), 556-563.

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

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

Plant, E. A., Hyde, J. S., Keltner, D., & Devine, P. G. (2000). The gender stereotyping of emotions. *Psychology of women quarterly*, *24*(1), 81-92.

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

Rauthmann, J. F., & Kolar, G. P. (2012). How "dark" are the Dark Triad traits? Examining the perceived darkness of narcissism, Machiavellianism, and psychopathy. *Personality and Individual Differences*, *53*(7), 884-889.

Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable facial configurations affect strategic choices in trust games with or without information about past behavior. *PloS one*, *7*(3), e34293.

Rogers, K. H., Le, M. T., Buckels, E. E., Kim, M., & Biesanz, J. C. (2018). Dispositional malevolence and impression formation: Dark Tetrad associations

with accuracy and positivity in first impressions. *Journal of personality*, *86*(6), 1050-1064.

Rowland, D. A., & Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE computer graphics and applications*, *15*(5), 70-76.

Rudert, S. C., Reutner, L., Greifeneder, R., & Walker, M. (2017). Faced with exclusion: Perceived facial warmth and competence influence moral judgments of social exclusion. *Journal of Experimental Social Psychology*, *68*, 101-112.

Rule, N. O., & Ambady, N. (2008). The face of success: Inferences from chief executive officers' appearance predict company profits. *Psychological science*, *19*(2), 109-111.

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

Sánchez-Villagra, M. R., & Van Schaik, C. P. (2019). Evaluating the self-domestication hypothesis of human evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, 28(3), 133-143.

Saribay, S. A., Biten, A. F., Meral, E. O., Aldan, P., Třebický, V., & Kleisner, K. (2018). The Bogazici face database: Standardized photographs of Turkish faces with supporting materials. *PloS one*, *13*(2), e0192018.

Schlicht, E. J., Shimojo, S., Camerer, C. F., Battaglia, P., & Nakayama, K. (2010). Human wagering behavior depends on opponents' faces. *PloS one*, *5*(7), e11663.

Shields, S. A., & Shields, S. A. (2002). *Speaking from the heart: Gender and the social meaning of emotion*. Cambridge University Press.

Shiramizu, V. K. M., Kozma, L., DeBruine, L. M., & Jones, B. C. (2019). Are dark triad cues really visible in faces?. *Personality and Individual Differences*, *139*, 214-216.

Spezio, M. L., Rangel, A., Alvarez, R. M., O'Doherty, J. P., Mattes, K., Todorov, A., ... & Adolphs, R. (2008). A neural basis for the effect of candidate appearance on election outcomes. *Social Cognitive and Affective Neuroscience*, *3*(4), 344-352.

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological science*, *21*(3), 349-354.

Sussman, A. B., Petkova, K., & Todorov, A. (2013). Competence ratings in US predict presidential election outcomes in Bulgaria. *Journal of Experimental Social Psychology*, *49*(4), 771-775.

Sutherland, C. A., Liu, X., Zhang, L., Chu, Y., Oldmeadow, J. A., & Young, A. W. (2018). Facial first impressions across culture: Data-driven modeling of Chinese and British perceivers' unconstrained facial impressions. *Personality and social psychology bulletin*, *44*(4), 521-537.

Sutherland, C. A., Rowley, L. E., Amoaku, U. T., Daguzan, E., Kidd-Rossiter, K. A., Maceviciute, U., & Young, A. W. (2015). Personality judgments from everyday images of faces. *Frontiers in Psychology*, *6*, 1616.

Tartaglia, S., & Rollero, C. (2015). The effects of attractiveness and status on personality evaluation. *Europe's journal of psychology*, *11*(4), 677.

Thielmann, I., & Hilbig, B. E. (2022). 17 Assumed similarity. *Cognitive Illusions: Intriguing Phenomena in Thinking, Judgment, and Memory*, 272.

Tingley, D. (2014). Face-off: Facial features and strategic choice. *Political Psychology*, *35*(1), 35-55.

Todorov, A., & Oh, D. (2021). The structure and perceptual basis of social judgments from faces. In *Advances in experimental social psychology* (Vol. 63, pp. 189-245). Academic Press.

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: a model based approach. *Social cognitive and affective neuroscience*, *3*(2), 119-127.

Todorov, A., Loehr, V., & Oosterhof, N. N. (2010). The obligatory nature of holistic processing of faces in social judgments. *Perception*, *39*(4), 514-532.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623-1626.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual review of psychology*, *66*, 519-545.

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, *27*(6), 813-833.

Tsankova, E., & Kappas, A. (2016). Facial skin smoothness as an indicator of perceived trustworthiness and related traits. *Perception*, *45*(4), 400-408.

Van't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition*, *108*(3), 796-803.

Walker, M., & Vetter, T. (2016). Changing the personality of a face: Perceived Big Two and Big Five personality factors modeled in real photographs. *Journal of personality and social psychology*, *110*(4), 609.

Walker, M., & Wänke, M. (2017). Caring or daring? Exploring the impact of facial masculinity/femininity and gender category information on first impressions. *PloS one*, *12*(10), e0181306.

Walker, M., Schönborn, S., Greifeneder, R., & Vetter, T. (2018). The Basel Face Database: A validated set of photographs reflecting systematic differences in Big Two and Big Five personality dimensions. *PloS one*, *13*(3), e0193190.

Wang, H., Hahn, A. C., DeBruine, L. M., & Jones, B. C. (2016). The motivational salience of faces is related to both their valence and dominance. *PloS one*, *11*(8), e0161114.

Wang, H., Han, C., Hahn, A. C., Fasolt, V., Morrison, D. K., Holzleitner, I. J., ... & Jones, B. C. (2019). A data-driven study of Chinese participants' social judgments of Chinese faces. *Plos one*, *14*(1), e0210315.

Wrangham, R. (2019). *The goodness paradox: how evolution made us both more and less violent.* Profile Books.

Widiger, T. A. (2009). Neuroticism. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 129–146). The Guilford Press.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological science*, *17*(7), 592-598.

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological science*, *26*(8), 1325-1331.

Zebrowitz, L. A. (1996). Physical Appearance as. *Stereotypes and stereotyping*, 79.

Zebrowitz, L. A. (1997). *Reading Faces: Window to the Soul?*, Boulder, Colorado.

Zebrowitz, L. A. (2017). First impressions from faces. *Current directions in psychological science*, *26*(3), 237-242.

Zebrowitz, L. A., & Collins, M. A. (1997). Accurate social perception at zero acquaintance: The affordances of a Gibsonian approach. *Personality and social psychology review*, *1*(3), 204-223.

Zebrowitz, L. A., & Montepare, J. (2006). The Ecological Approach to Person Perception: Evolutionary Roots and Contemporary Offshoots.

Zebrowitz, L. A., & Montepare, J. M. (1992). Impressions of babyfaced individuals across the life span. *Developmental psychology*, *28*(6), 1143.

Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and personality psychology compass*, *2*(3), 1497-1517.

Zebrowitz, L. A., Fellous, J. M., Mignault, A., & Andreoletti, C. (2003). Trait impressions as overgeneralized responses to adaptively significant facial qualities: Evidence from connectionist modeling. *Personality and social psychology review*, *7*(3), 194-215.

Zebrowitz, L. A., Wang, R., Bronstad, P. M., Eisenberg, D., Undurraga, E., Reyes-García, V., & Godoy, R. (2012). First impressions from faces among US and culturally isolated Tsimane'people in the Bolivian rainforest. *Journal of Cross-Cultural Psychology*, *43*(1), 119-134.

Zanella, M., Vitriolo, A., Andirko, A., Martins, P. T., Sturm, S., O'Rourke, T., ... & Testa, G. (2019). Dosage analysis of the 7q11. 23 Williams region identifies BAZ1B as a major human gene patterning the modern human face and underlying self-domestication. *Science advances*, 5(12), eaaw7908.

APPENDICES

APPENDIX A: Paired Samples T-test Results of Pilot Study

The results of the paired samples t-test conducted for each manipulated face pair. 'BFD No' refers to the number assigned to each face identity in the Bogazici Face Database. The numbers in bold belong to the face pairs used in the main study.

		BFD No	Pair	Mean	SD	SE Mean	95% CI Lower	95% CI Upper	t	df	Sig. (2-tailed)	Cohen's d
Skin Smoothness	**Female**	155	High vs. Low	5,43	2,31	0,51	4,38	6,48	10.748	20	.000	3,33
		156	**High vs. Low**	**5,43**	**1,72**	**0,38**	**4,65**	**6,21**	**14.466**	**20**	**.000**	4,09
		176	High vs. Low	5,10	2,32	0,51	4,04	6,15	10.057	20	.000	3,08
		183	High vs. Low	5,52	1,97	0,43	4,63	6,42	12.881	20	.000	4,20
		188	High vs. Low	5,62	1,66	0,36	4,86	6,37	15.534	20	.000	4,22
		206	High vs. Low	5,14	1,82	0,40	4,31	5,97	12.918	20	.000	3,66
		227	High vs. Low	5,57	1,99	0,43	4,67	6,48	12.835	20	.000	3,77
		232	High vs. Low	5,81	1,50	0,33	5,12	6,49	17.702	20	.000	5,46
	Male	**103**	**High vs. Low**	6,14	2,35	0,51	5,07	7,21	11.972	20	.000	4,52
		119	High vs. Low	5,38	1,96	0,43	4,49	6,27	12.571	20	.000	3,74
		122	High vs. Low	5,05	2,25	0,49	4,02	6,07	10.296	20	.000	3,15
		141	High vs. Low	4,62	2,89	0,63	3,30	5,93	7.326	20	.000	2,52
		161	High vs. Low	4,05	1,99	0,43	3,14	4,95	9.336	20	.000	1,96
		175	High vs. Low	4,24	2,70	0,59	3,01	5,47	7.193	20	.000	1,91
		224	High vs. Low	4,86	1,98	0,43	3,95	5,76	11.230	20	.000	2,92
		240	High vs. Low	4,67	2,13	0,46	3,70	5,64	10.044	20	.000	2,41
Masculinity	**Female**	380	High vs. Low	3,81	1,54	0,34	3,11	4,51	11.359	20	.000	2,96
		398	High vs. Low	3,90	1,45	0,32	3,25	4,56	12.376	20	.000	2,88
		429	High vs. Low	3,57	2,20	0,48	2,57	4,57	7.426	20	.000	2,32
		454	**High vs. Low**	4,62	1,66	0,36	3,86	5,37	12.770	20	.000	3,64
		460	High vs. Low	4,29	1,82	0,40	3,46	5,11	10.788	20	.000	3,12
		473	High vs. Low	3,24	3,14	0,69	1,81	4,67	4.718	20	.000	1,72
		474	High vs. Low	3,62	2,31	0,50	2,57	4,67	7.172	20	.000	2,55
		492	High vs. Low	4,00	1,79	0,39	3,19	4,81	10.247	20	.000	3,16
	Male	**319**	**High vs. Low**	3,90	2,23	0,49	2,89	4,92	8.010	20	.000	2,20
		322	High vs. Low	2,33	2,46	0,54	1,22	3,45	4.353	20	.000	1,19
		327	High vs. Low	2,81	1,54	0,34	2,11	3,51	8.377	20	.000	1,61
		337	High vs. Low	2,62	1,94	0,42	1,74	3,50	6.200	20	.000	1,45
		356	High vs. Low	3,00	2,55	0,56	1,84	4,16	5.392	20	.000	1,57
		358	High vs. Low	2,57	2,01	0,44	1,65	3,49	5.850	20	.000	1,59
		372	High vs. Low	2,86	2,33	0,51	1,80	3,92	5.620	20	.000	1,81
		401	High vs. Low	2,90	2,14	0,47	1,93	3,88	6.213	20	.00	1,92

		BFD No	Pair	Mean	SD	SE Mean	95% CI Lower	95% CI Upper	t	df	Sig. (2-tailed)	Cohen's d
Baby facedness	**Female**	**238**	**High vs. Low**	5,14	1,65	0,36	4,39	5,89	14.267	20	.000	4,08
		258	High vs. Low	3,71	2,65	0,58	2,51	4,92	6.427	20	.000	1,92
		302	High vs. Low	3,86	2,10	0,46	2,90	4,82	8.399	20	.000	2,06
		334	High vs. Low	2,81	2,96	0,65	1,46	4,16	4.350	20	.000	1,38
		340	High vs. Low	3,67	2,18	0,47	2,68	4,66	7.723	20	.000	2,01
		355	High vs. Low	3,43	2,98	0,65	2,07	4,78	5.279	20	.000	1,76
		366	High vs. Low	3,62	2,16	0,47	2,64	4,60	7.693	20	.000	1,94
		378	High vs. Low	3,90	2,19	0,48	2,91	4,90	8.176	20	.000	2,38
	Male	241	High vs. Low	3,10	2,59	0,56	1,92	4,27	5.484	20	.000	1,62
		247	**High vs. Low**	3,33	2,13	0,46	2,36	4,30	7.174	20	.000	1,75
		252	High vs. Low	2,95	2,29	0,50	1,91	4,00	5.906	20	.000	1,73
		268	High vs. Low	2,19	2,75	0,60	0,94	3,44	3.650	20	.002	1,06
		289	High vs. Low	3,33	2,35	0,51	2,26	4,40	6.494	20	.000	1,60
		293	High vs. Low	2,90	2,47	0,54	1,78	4,03	5.394	20	.000	1,55
		313	High vs. Low	3,76	2,17	0,47	2,78	4,75	7.960	20	.000	2,00
		317	High vs. Low	3,48	1,83	0,40	2,64	4,31	8.688	20	.000	1,95

APPENDIX B: Paired Samples T-test Results Based On Response Time

The results of the paired samples t-tests conducted to assess whether faces with high and low levels of skin smoothness (S), baby facedness (B) and masculinity (MA) differed in their trait ratings, separated with respect to overall response time (see table below). Namely, in the 'Original' row, results of the main analyses are reported in which all participants who took part in the study (N = 505) were included in the analyses, irrespective of response time. In the '7 mins' row, only participants who finished the study in over 7 minutes (N =478) were included in the analyses, those who finished under 7 minutes were excluded. In the '10 mins' row, only participants who finished the study in over 9 minutes (N =393) were included in the analyses, those who finished under 9 minutes were excluded.

As can be seen in the table provided below, when comparing the three samples, no meaningful differences emerged.

		Pair	Mean	*SD*	*SE* Mean	95% CI Lower	95% CI Upper	t	df	$P_{bonferroni}$	Cohen's d
Psychopathy	**Original**	S_High_ - S_Low	0,35	1,65	0,07	0,20	0,49	4,75	504	<.001	0.21
		MA_High -MA_Low	1,03	1,95	0,09	0,86	1,20	11,94	504	<.001	0.53
		B_High - B_Low	-1,01	1,92	0,09	-1,18	-0,84	-11,81	504	<.001	0.53
	7 mins	S_High_ - S_Low	0,33	1,64	0,07	0,18	0,47	4,35	478	<.001	0.19
		MA_High -MA_Low	1,05	1,94	0,09	0,88	1,23	11,88	478	<.001	0.65
		B_High - B_Low	-0,99	1,92	0,09	-1,16	-0,81	-11,26	478	<.001	0.60
	9 mins	S_High_ - S_Low	0,28	1,62	0,08	0,12	0,44	3,41	393	.011	0.17
		MA_High -MA_Low	1,07	1,98	0,10	0,87	1,27	10,75	393	<.001	0.66
		B_High - B_Low	-0,98	1,95	0,10	-1,18	-0,79	-10,01	393	<.001	0.60
Machiavellianism	**Original**	S_High_ - S_Low	0,43	1,71	0,08	0,28	0,58	5,63	504	<.001	0.32
		MA_High -MA_Low	0,82	2,01	0,09	0,64	0,99	9,14	504	<.001	0.46
		B_High - B_Low	-0,87	1,97	0,09	-1,04	-0,69	-9,88	504	<.001	0.44
	7 mins	S_High_ - S_Low	0,42	1,71	0,08	0,27	0,57	5,37	478	<.001	0.25
		MA_High -MA_Low	0,83	2,00	0,09	0,65	1,01	9,01	478	<.001	0.50
		B_High - B_Low	-0,84	1,95	0,09	-1,02	-0,67	-9,44	478	<.001	0.51
	9 mins	S_High_ - S_Low	0,37	1,70	0,09	0,20	0,54	4,29	393	<.001	0.22
		MA_High -MA_Low	0,90	2,04	0,10	0,70	1,10	8,72	393	<.001	0.55
		B_High - B_Low	-0,86	1,92	0,10	-1,05	-0,67	-8,91	393	<.001	0.52
Narcissism	**Original**	S_High_ - S_Low	0,52	1,65	0,07	0,38	0,67	7,14	504	<.001	0.25
		MA_High -MA_Low	0,94	2,05	0,09	0,76	1,12	10,30	504	<.001	0.41
		B_High - B_Low	-0,86	1,96	0,09	-1,04	-0,69	-9,89	504	<.001	0.44
	7 mins	S_High_ - S_Low	0,52	1,65	0,08	0,37	0,67	6,88	478	<.001	0.32
		MA_High -MA_Low	0,96	2,07	0,09	0,78	1,15	10,21	478	<.001	0.59
		B_High - B_Low	-0,84	1,98	0,09	-1,01	-0,66	-9,22	478	<.001	0.50
	9 mins	S_High_ - S_Low	0,53	1,64	0,08	0,37	0,69	6,43	393	<.001	0.33
		MA_High -MA_Low	1,04	2,06	0,10	0,83	1,24	10,00	393	<.001	0.64
		B_High - B_Low	-0,81	1,98	0,10	-1,01	-0,62	-8,17	393	<.001	0.48
Neuroticism	**Original**	S_High_ - S_Low	-0,03	1,57	0,07	-0,17	0,11	-0,41	504	1	0.02
		MA_High -MA_Low	0,18	1,91	0,09	0,02	0,35	2,17	504	.463	0.10
		B_High - B_Low	-0,14	1,73	0,08	-0,29	0,01	-1,82	504	1	0.08
	7 mins	S_High_ - S_Low	-0,02	1,55	0,07	-0,16	0,12	-0,28	478	1	0.01
		MA_High -MA_Low	0,20	1,92	0,09	0,02	0,37	2,23	478	.393	0.10
		B_High - B_Low	-0,12	1,72	0,08	-0,28	0,03	-1,55	478	1	0.08
	9 mins	S_High_ - S_Low	-0,05	1,55	0,08	-0,20	0,11	-0,58	393	1	0.03
		MA_High -MA_Low	0,22	1,93	0,10	0,03	0,42	2,31	393	.323	0.12
		B_High - B_Low	-0,13	1,74	0,09	-0,30	0,05	-1,43	393	1	0.08

		Pair	Mean	SD	SE Mean	95% CI Lower	95% CI Upper	t	df	$P_{bonferroni}$	Cohen's d
Openness	**Original**	S_High_ - S_Low	0,02	1,46	0,06	-0,11	0,15	0,27	504	1	0.01
		MA_High -MA_Low	-0,49	1,74	0,08	-0,65	-0,34	-6,35	504	<.001	0.28
		B_High - B_Low	0,55	1,78	0,08	0,40	0,71	6,98	504	<.001	0.31
	7 mins	S_High_ - S_Low	0,02	1,41	0,06	-0,11	0,14	0,24	478	1	0.01
		MA_High -MA_Low	-0,53	1,75	0,08	-0,69	-0,37	-6,61	478	<.001	0.35
		B_High - B_Low	0,58	1,77	0,08	0,42	0,74	7,12	478	<.001	0.39
	9 mins	S_High_ - S_Low	-0,01	1,43	0,07	-0,15	0,14	-0,07	393	1	0.00
		MA_High -MA_Low	-0,58	1,77	0,09	-0,75	-0,40	-6,46	393	<.001	0.38
		B_High - B_Low	0,59	1,79	0,09	0,42	0,77	6,58	393	<.001	0.39
Extraversion	**Original**	S_High_ - S_Low	0,30	1,63	0,07	0,16	0,44	4,14	504	<.001	0.18
		MA_High -MA_Low	-0,30	1,68	0,07	-0,45	-0,15	-4,00	504	.001	0.18
		B_High - B_Low	0,29	1,85	0,08	0,13	0,45	3,54	504	.007	0.16
	7 mins	S_High_ - S_Low	0,27	1,65	0,08	0,12	0,42	3,58	478	.006	0.17
		MA_High -MA_Low	-0,30	1,71	0,08	-0,46	-0,15	-3,90	478	.002	0.20
		B_High - B_Low	0,29	1,86	0,09	0,12	0,46	3,41	478	.010	0.20
	9 mins	S_High_ - S_Low	0,26	1,63	0,08	0,10	0,42	3,20	393	.023	0.17
		MA_High -MA_Low	-0,29	1,73	0,09	-0,47	-0,12	-3,38	393	.012	0.19
		B_High - B_Low	0,31	1,85	0,09	0,13	0,49	3,34	393	.014	0.20
Agreeableness	**Original**	S_High_ - S_Low	-0,15	1,62	0,07	-0,29	-0,01	-2,08	504	.568	0.09
		MA_High -MA_Low	-0,99	1,80	0,08	-1,15	-0,83	-12,39	504	<.001	0.55
		B_High - B_Low	0,82	1,92	0,09	0,65	0,99	9,59	504	<.001	0.43
	7 mins	S_High_ - S_Low	-0,13	1,63	0,07	-0,28	0,01	-1,79	478	1	0.09
		MA_High -MA_Low	-1,03	1,80	0,08	-1,19	-0,87	-12,50	478	<.001	0.71
		B_High - B_Low	0,79	1,93	0,09	0,62	0,96	8,97	478	<.001	0.53
	9 mins	S_High_ - S_Low	-0,10	1,64	0,08	-0,26	0,06	-1,23	393	1	0.08
		MA_High -MA_Low	-1,08	1,81	0,09	-1,26	-0,90	-11,80	393	<.001	0.75
		B_High - B_Low	0,84	1,91	0,10	0,65	1,03	8,72	393	<.001	0.57
Conscientiousness	**Original**	S_High_ - S_Low	0,04	1,51	0,07	-0,10	0,17	0,53	504	1	0.02
		MA_High -MA_Low	-0,99	1,69	0,08	-1,14	-0,84	-13,15	504	<.001	0.59
		B_High - B_Low	-0,09	1,77	0,08	-0,25	0,06	-1,19	504	1	0.05
	7 mins	S_High_ - S_Low	0,04	1,50	0,07	-0,10	0,17	0,54	478	1	0.03
		MA_High -MA_Low	-1,02	1,70	0,08	-1,17	-0,86	-13,07	478	<.001	0.70
		B_High - B_Low	-0,13	1,79	0,08	-0,29	0,03	-1,60	478	1	0.09
	9 mins	S_High_ - S_Low	0,03	1,49	0,08	-0,12	0,18	0,41	393	1	0.02
		MA_High -MA_Low	-1,01	1,67	0,08	-1,18	-0,84	-12,00	393	<.001	0.71
		B_High - B_Low	-0,10	1,79	0,09	-0,28	0,08	-1,08	393	1	0.06

Note. Cohen's d values are reported as absolute values.