PERCEPTIONS OF INTELLIGENCE AND THE ATTRACTIVENESS HALO

Sean N. Talamas

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

2016

Full metadata for this item is available in St Andrews Research Repository at:
http://research-repository.st-andrews.ac.uk/

Please use this identifier to cite or link to this item:
http://hdl.handle.net/10023/10851

This item is protected by original copyright
Perceptions of Intelligence and the 
Attractiveness Halo

Sean N. Talamas

University of St Andrews

This thesis is submitted in partial fulfilment for the degree of

PhD

at the

University of St Andrews

Date of Submission

November 18th, 2015
Candidate’s Declarations:

I, Sean N. Talamas, hereby certify that this thesis, which is approximately 37,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in January 2014 and as a candidate for the degree of PhD in January 2014 (direct entry into second year of PhD studies); the higher study for which this is a record was carried out in the University of St Andrews between January 2014 and November 2015.

Date __________________ Signature of Candidate __________________________

Supervisor’s Declaration:

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

Date_____________________ Signature of Supervisor________________________
Permission for publication

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

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

PRINTED AND ELECTRONIC COPY

Embargo on all or part of print and electronic copy for a period of 2 years on the following ground(s):

- Publication would preclude future publication

Supporting statement for printed and electronic embargo request:

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

Date _______________

Signature of Candidate ______________________

Signature of Supervisor ______________________
Collaboration Statement

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

This thesis is based partly on works submitted to publication in peer-reviewed academic journals. Co-authors are listed at the beginning of each empirical chapter (Chapters 4-6) in which they are featured. While co-authors are those that contributed intellectually to the presented work, many others also contributed to works presented in this thesis. Anne Perrett, Kate Epstein, and Iris Holzleitner provided assistance in proof reading. Daniel Menendez, Charlotte Poynton, Stephanie Belenkov, Louise Milligan, Emma Martin, Rebecca Munns, Natalia Fedorova, and Olivia Ives assisted in face delineation and data collection. Dengke Xiao provided the experimental interface for face rating tasks and cognitive ability tests, as well as the interface used to objectively measure facial cues.
Acknowledgments

This PhD is far from my own work. I could not have summoned the effort, motivation, or intellectual curiosity the project required alone. Indeed, left to myself, I can be unmotivated and overwhelmed by my scattered curiosities. My intellectual ability is unreliable, to say the least.

Then, for instilling in me the necessary attributes the project required, I give thanks to my Grandmother Nati Talamas and Uncle Joe Talamas for their mentorship and guidance. Words will never be able to express my gratitude. Any and all good I have ever done, or will ever do, I owe to lessons they imparted to me. My passion for teaching and learning is undeniably inspired by their encouragement and the passion both of them have for these pursuits.

The US Military (TACP community) developed in me some of the qualities and provided the funding needed to complete this PhD. Indeed the military made starting the PhD possible. I thank TACP leadership for letting me pursue my education while deployed and a Masters afterward, rather than another deployment.

Thank you to the University of St Andrews rowing and basketball team for assistance with participant recruitment – cheers for not blocking me on Facebook after the first (and second) request. A special shout out to the basketball team for consistently dragging me out of the office to train, play, or coach. Messing about with you guys kept me sane. Many thanks to Chris P, Andreas P, and Dan K, for their reliable assistance off court, which was always better than my assists on court.

I am particularly indebted to the Perception Lab for providing feedback on my research ideas, concerns, posters, manuscripts, and presentations, taking care of Willypete (and getting full of his hair), and for putting up with my often-immature humor. The diverse group of individuals at the Perception Lab offered me additional outlets for this humor and other means of staying sane during the process: Audrey’s Cards Against Humanity skills have made me feel less guilty about my own inappropriate humor; Jing’s dance skills have emphasized the importance of a work-life balance; Martha’s neon posters have prompted my life-planning skills; Steph’s willingness to attend to Carlotta has encouraged my patience; and Carlotta’s gossiping skills have resulted in surprisingly
decent advice on multiple occasions. Thank you also to Lotty, a wannabe perception lab member/wife who, like the rest of the perception lab, has always had my back. Honestly, you all are amazing—I thank you so much for everything.

A special shout-out to Iris who took a considerable amount of time out of her PhD work to help get me settled into the lab on my first day and every day after that for the next two years. A big thanks to the various past Perception Lab members like Ross W, Amanda H, Lisa D, Ben J, Tony L, and Ian P, for conversations that sparked ideas for improved and new research. Also, thank you Lesley for your unfailing administrative support, Dengke for assistance with multiple technical queries, and Anne, for proof reading, which has made me a better writer.

To my second supervisor (my only supervisor with natural-colored hair) Martin Campbell for providing constructive feedback on all my work in an impressively timely and reliable fashion and my co-supervisor Ken Mavor for invaluable contributions to manuscripts and to my understanding of statistics.

To my boy Willypete (A.K.A. William Peter) for his assistance with participant (and friend) recruitment. I have benefited from your St. Andrews fame and appreciate your moral support throughout.

Finally, I will be forever appreciative of my supervisor, Dave Perrett. Dave’s mentorship far surpassed his role as a supervisor: his daily support was unmatched, his patience was extraordinary, his passion for research was contagious, and his concern for my work-life balance was sobering. Thank you, Dave.
# Table of Contents

**Thesis Abstract** .................................................................................................................. 1

**Overview of Introductory Chapters** .................................................................................. 3

**Chapter 1: What Is Attractiveness?** .................................................................................. 5
  1. Attractiveness .................................................................................................................. 7
     1.1 Symmetry .................................................................................................................. 8
     1.2 Averageness .............................................................................................................. 10
     1.3 Sexual Dimorphism ................................................................................................. 11
     1.4 Skin Condition ......................................................................................................... 13
     1.5 Facial Adiposity ....................................................................................................... 15
     1.6 Interacting Cues ...................................................................................................... 15
     1.7 Stereotypes and Overgeneralizations ..................................................................... 16
     1.8 Summary .................................................................................................................. 19

**Chapter 2: What Is Intelligence?** .................................................................................... 21
  2. Intelligence ..................................................................................................................... 23
     2.1 Theories of Intelligence ......................................................................................... 23
     2.2 Outcomes of Intelligence ....................................................................................... 27
     2.3 Summary .................................................................................................................. 29

**Chapter 3: Attractiveness-Intelligence Halo** ................................................................. 31
  3. Halo .................................................................................................................................. 33
     3.1 Attractiveness and Intelligence Halo ...................................................................... 33
     3.2 Attractiveness and the Expectancy Effect .............................................................. 35
     3.3 Intelligence and Attractiveness Halo ..................................................................... 37
     3.4 Individual Differences ............................................................................................ 40
     3.5 Summary .................................................................................................................. 42

**Overview of Experimental Chapters** ............................................................................ 44

**Chapter 4: Perceived Intelligence: Beyond the Attractiveness Halo** ......................... 47
  4. Abstract .......................................................................................................................... 49
     4.1 Introduction .............................................................................................................. 51
     4.2 Overview of the Studies .......................................................................................... 55
     4.3 Study 1: Adult Faces ............................................................................................... 57
     4.4 Study 2: Children’s Faces ....................................................................................... 75
     4.5 Study 3: Malleable Facial Cues Transform ............................................................. 83
     4.6 Study 4: Sleep-Restricted Faces ............................................................................. 87
List of Tables and Figures

Figure 4.1. Eyelid-openness calculation. 72
Figure 4.2. Mouth curvature measure. 73
Figure 4.3. Averaged adult images. 78
Table 4.1. Zero-order correlation matrix for adult stimuli. 80
Table 4.2. Direct and mediated effects of eyelid-openness, mouth curvature, and gender on perceived attractiveness and perceived intelligence. 82
Figure 4.4. Linear regression model for adult faces. 82
Table 4.3. Direct and mediated effects of eye size, mouth curvature, and gender on perceived attractiveness and perceived intelligence. 83
Table 4.4. Direct and mediated effects of eye size, eyelid-openness, mouth curvature, and gender on perceived attractiveness and perceived intelligence. 84
Figure 4.5. Averaged children images. 92
Table 4.5. Zero-order correlation matrix for children faces. 93
Table 4.6. Direct and mediated effects of eyelid-openness, mouth curvature, and gender on perceived attractiveness and perceived intelligence of children’s faces 94
Figure 4.6. Linear regression model of the perception of children’s faces. 94
Figure 4.7. Malleable facial cues transform. 98
Figure 4.8. Averaged sleep-restriction images. 106
Table 4.7. Direct and mediated effects of change in eyelid-openness, change in mouth curvature, and gender on changes in perceived attractiveness and perceived intelligence of adult faces with and without sleep restriction. 109
Figure 4.9. Linear regression model of the perceptual effects of sleep restriction. 109
Figure 5.1. Variation in strength of the individual attractiveness-intelligence halo across individuals. 128
Table 5.1. Zero-order correlation matrix 130
Figure 5.2. Participants’ actual intelligence and halo scatter plot. 131

Table 5.2. Hierarchal linear regression 132

Table 6.1. Sample information 151

Figure 6.1. Averaged perceived conscientiousness. 153

Table 6.2. Zero-order correlation matrix 154

Figure 6.2. Partial correlation bar graph. 155

Figure 6.3. Suppression effect. 157
Thesis Abstract

Perceptions of intelligence are strongly related to attractiveness and have a significant impact on first impressions. The introductory chapters (1–3) provide an overview of the literature on attractiveness, halo effects, and intelligence, while the experimental chapters (4–6) explore perceptions of cues to intelligence beyond attractiveness, individual differences in the susceptibility to the halo, and the accuracy of perceptions of competence.

Chapter 4 investigated the malleable facial cues of eyelid-openness and mouth curvature and their influence on perceived intelligence. Attractiveness partially mediated intelligence impression, but effects of eyelid-openness and subtle smiling enhanced intelligence ratings independent of attractiveness. These effects were observed and replicated in between individual (cross-sectional) studies of natural images of adult faces, child faces, through digital manipulation of individual cues in the same faces, and in a within individual sleep-restricted sample. Chapter 5 investigated the relationship between perceived intelligence and attractiveness by exploring whether a raters own intelligence may be related to a stronger endorsement of the perceived intelligence-attractiveness halo. The correlation between ratings of the perceived intelligence and attractiveness was found to be stronger for participants who scored higher on an intelligence test than participants with lower intelligence test scores. Chapter 6 investigated the limiting effects of attractiveness on perceptions of competence. When statistically controlling for the attractiveness halo, academic performance could be predicted from judgments of conscientiousness but not from ratings of intelligence.

Thus this thesis demonstrates that malleable facial cues can influence perceptions of intelligence independent of attractiveness, identifies an individual difference that influences endorsement of the intelligence-attractiveness halo, and shows the limiting effects of the attractiveness halo on potentially accurate perceptions of academic performance. Collectively these findings provide evidence of the powerful influence of attractiveness on perceptions of intelligence; such work is necessary if we are to mitigate such bias.
Overview of Introductory Chapters

The focus of this thesis will be on perceptions of intelligence. It will take into account the robust influence of facial attractiveness. The introductory chapters (Chapters 1–3) discuss relevant literature on facial attractiveness, theories of intelligence, and the halo effect. These chapters will provide a broad overview, while the introduction to the experimental chapters (Chapters 4–6) will provide literature more specific to the research questions investigated.

Chapter 1 summarizes literature on fixed and more changeable cues to facial attractiveness, including facial symmetry, averageness, sexual dimorphism, skin condition, and adiposity. Given the strong inter-correlations between attractiveness, health, and intelligence (arguably related to a general "fitness factor"), it is likely that the cues discussed could influence perceived intelligence in a similar way to their influence on perceived attractiveness. Chapter 4 investigates the role of additional malleable facial cues that may influence perceived intelligence and attractiveness given their relationship to alertness and perceived health.

Chapter 2 considers the different theories and definitions of intelligence and provides a brief overview of the relationship between intelligence and resource acquisition, mate selection, and academic achievement. Research shows academic performance correlates strongly with intelligence and even more strongly with personality. Thus, Chapter 6 explores whether perceptions of conscientiousness (rather than perceived intelligence specifically) will improve accuracy of perceived academic performance when controlling for an attractiveness bias.

Chapter 3 further explores the relationship between perceived intelligence and attractiveness by scrutinizing expectancy and general halo effects. Chapter 5 explores whether individual differences in intelligence affect the endorsement of the perceived intelligence-attractiveness halo.
Chapter 1: What Is Attractiveness?
1. Attractiveness

Physical attractiveness has long been associated with a favorable bias. Dion, Berscheid, and Walster (1972) claimed that “what is beautiful is good” (p.285). Since that time, research has consistently shown that physical attractiveness has a halo effect, leading to consistent, but not necessarily accurate, positive personal quality judgments of physically attractive people. This chapter identifies the facial cues that people consider to be attractive.

Tradition holds that beauty is in the eye of the beholder. Some research supports this claim, finding that differences in the beholder (discussed more in Chapter 3) include hormones (DeBruine, Jones, & Perrett, 2005; Jones et al., 2008; Penton-Voak et al., 1999; Penton-Voak & Perrett, 2000), personality (Little, Burt, & Perrett, 2006), and self-evaluation (Kandrik & DeBruine, 2012; Little, Burt, Penton-Voak, & Perrett, 2001), all of which are just some individual differences that can influence mate preferences (see Germine et al., 2015; Perrett, 2010). Despite these differences in views on beauty, research has also found a surprising degree of agreement regarding facial attractiveness. Specifically, cross-cultural studies have found high levels of agreement on facial attractiveness across different countries (Langlois et al., 2000; Swami, Knight, Tovée, Davies, & Furnham, 2007; Swami & Tovée, 2005; Tovée, Swami, Furnham, & Mangalparsad, 2006) and across gender (Rhodes, 2006).

It is possible that agreement over facial attractiveness stems from cues to attractiveness that are related to actual health. In a comprehensive review on the evolutionary psychology of facial beauty, Rhodes (2006) discusses a “good genes” theory of mate selection, which argues that the fixed facial cues of averageness, symmetry, and sexual dimorphism are all good candidates for biologically based standards of beauty, given their relation to health (Fink & Penton-Voak, 2002; Grammer, Fink, Möller, & Thornhill, 2003; Johnston & Franklin, 1993; Scheib, Gangestad, & Thornhill, 1999; Thornhill & Gangestad, 1993). More recently, research has explored the influence of the malleable facial cues of skin condition and adiposity in relation to attractiveness and actual health (Coetzee, Chen, Perrett, & Stephen, 2010; Coetzee, Perrett, & Stephen, 2009; Rantala et al., 2013; Stephen, Coetzee, Law Smith, & Perrett, 2009; Stephen, Coetzee, &
Measurement of these cues is of particular importance given the subtlety of cues like symmetry and their relationship to perceived and actual attributes. Building on this research, this chapter discusses the fixed facial cues of symmetry, averageness, and sexual dimorphism, as well as malleable cues of skin condition and adiposity and their relationship to attractiveness and health.

Stereotypes and overgeneralizations are likely to influence impression formation. For instance, the perception of individuals with droopy eyelids as tired and unintelligent may be an overgeneralized response based on the well-known relationship between sleep and decreased cognitive performance (Pilcher & Huffcutt, 1996; Thomas et al., 2000). An overview of stereotypes, the overgeneralization effect, and valid trait attributions from faces serves as an introduction to the new malleable cues (eyelid-openness and mouth curvature) that are investigated in relation to perceived intelligence in Chapter 4.

### 1.1 Symmetry

Research suggests that symmetry is a signal of high-quality genes and relative disease resistance (Perrett et al., 1999; Rhodes et al., 2001; Scheib et al., 1999; Thornhill & Gangestad, 1993). The methods used to measure the effects of symmetry are particularly important, given that the differences in methods has led to differences in findings (these methods are also important when considering the cues investigated in Chapter 4). Earlier studies on facial symmetry assessed preferences for symmetry by producing chimeras, which are made by taking one vertical half of a face and mirroring it (split by a vertical midline) to create the other half (Kowner, 1996; Samuels, Butterworth, Roberts, Graupner, & Hole, 1994). This method can produce symmetric, but distorted faces with non-average features which were found to be less attractive than asymmetric faces (Perrett et al., 1999).

Studies of un-manipulated faces have found that participants consider fluctuating asymmetries (FAs) unattractive. While directional asymmetries, like the mammalian heart for example, reflect a systematic difference in one side of a symmetrical plane, FAs are random deviations from symmetry that have been associated with challenges to fetal growth during pregnancy (Blount et al., 1997), or early post-birth development (e.g., exposures to toxins or high levels of stress; Thornhill & Gangestad, 1996, 1999). Humans, as well as other species, are able to detect small FAs, without any training (Palmer &
In turn, measures of symmetry have shown to correlate with facial attractiveness (Jones, DeBruine, & Little, 2007; Penton-Voak et al., 2001; Perrett et al., 1999).

Several studies have suggested a relationship between symmetry and health, whether self-reported (Thornhill & Gangestad, 2006) or actual (Jones et al., 2004; Rhodes et al., 2001). FAs increase with age (Wilson & Manning, 1996) and have shown to correlate with heavier body weight (Manning, 1995), predisposition to breast cancer (Manning, Scutt, Whitehouse, & Leinster, 1997), slower running speed (Manning & Pickup, 1998), and higher metabolic rate (Manning, Koukourakis, & Brodie, 1997).

Similarly, studies have consistently shown a relationship between FAs and intelligence (Bates, 2007; Furlow, Armijo-Prewitt, Gangestad, & Thornhill, 1997; Luxen & Buunk, 2006; Penke et al., 2009). A meta-analysis by Banks, Batchelor, and McDaniel (2010) based on 14 samples across 1871 people found that people who score higher on intelligence tests are significantly more symmetrical, although this correlation is weak (i.e., effect size is small).

The relationship between intelligence and symmetry is one indicator of a "general fitness factor", in which multiple cues are convergent indicators of survival and reproductive success (Furlow et al., 1997; Miller, 2000; Möller & Alatalo, 1999; Prokosch, Coss, Scheib, & Blozis, 2009; Prokosch, Yeo, & Miller, 2005; Singh, 1995). Intelligence has also been shown to correlate with physical health (Arden, Gottfredson, & Miller, 2009; Batty, Deary, & Gottfredson, 2007; Gottfredson & Deary, 2004; but see Pound et al., 2014), semen quality (Arden, Gottfredson, Miller, & Pierce, 2009), and life expectancy (Jokela, Batty, Deary, Gale, & Kivimäki, 2009). Hence, the relationship between intelligence and a general "fitness factor" highlights the potential for visible cues to intelligence to be perceived as attractive (Moore, Filippou, & Perrett, 2011).

In support of facial symmetry being related to gene quality and attractiveness, Little, Burt, Penton-Voak, and Perrett (2001) found that women who rated themselves as attractive preferred symmetrical male faces, while women who rated themselves as average or below average had lower standards for symmetrical faces in males. Additionally, Little and Jones (2012) found that women in different phases of their menstrual cycles have different preferences for symmetry. Specifically, women were found to prefer symmetry in
the follicular phase of their cycle when fertility is likely to be highest compared to the luteal (premenstrual) phase (Little & Jones, 2012). This effect was observed when rating faces for short-term relationship attractiveness, but no influence of cycle was seen for long-term judgments, consistent with previous literature on mate preferences across the ovulatory cycle (Gildersleeve et al., 2013; Gildersleeve, Haselton, & Fales, 2014). Collectively, stronger preferences for symmetry by attractive and fertile women provides support for symmetry being a signal for genetic quality and being a desired trait.

1.2 Averageness

Research on facial averaging has come a long way from when Galton (1878) first noted that a composite of faces is better looking than individual faces. Several methods of measuring and assessing the attractiveness of averaged faces have developed since Galton’s initial observation. While Langlois and Roggman (1990) computed mathematical averages of digital images and argued that averaged faces were more attractive because they are more similar to the average value of the population, these methods and theories have advanced since then.

Researchers criticizing studies involving computer-generated average faces have argued that participants respond favorably to homogenous skin texture and symmetry rather than averageness itself (Alley & Cunningham, 1991; Benson & Perrett, 1992; Tiddeman, Burt, & Perrett, 2001). Yet, while Grammer and Thornhill (1994) found that averaged faces were not attractive when statistically controlling for the influence of symmetry, research by Rhodes, Sumich, and Byatt (1999) found that actually symmetrical faces became more attractive when averageness was increased. Additionally, more recent research by Jones, Debruine, and Little (2007) found average faces were more attractive than less average faces when symmetry was digitally controlled for in composite faces.

The attractiveness of averageness has important caveats. Averageness is attractive within categories like age, sex, and race (Bestelmeyer et al., 2008; Little, DeBruine, Jones, & Waitt, 2008), but averages made between categories are not necessarily more attractive. For example, an average of women and men may look less attractive than an average of just women because femininity is attractive in women’s faces (Perrett et al., 1998). Averaging also has limitations with respect to more attractive faces (DeBruine, Jones,
Unger, Little, & Feinberg, 2007). A set of faces deemed attractive will produce a more attractive average by itself than the average of a larger set of faces that includes all of the attractive faces and a set of less attractive faces (Perrett, May, & Yoshikawa, 1994).

The notion that an average face is more attractive than a unique face may run contrary to popular belief. Yet, it can be argued that unique faces are categorized as unusual, different, or weird which may be subtly related to illness or developmental abnormalities. Correspondingly, averageness in faces can hint at probable good health, much like symmetry does. Theory would suggest that traits that advertise resistance to parasites (e.g., bacteria and viruses) are preferred (Grammer & Thornhill, 1994; Hamilton & Zuk, 1982). Indeed, research has found that both facial averageness and symmetry were perceived as healthier (Fink, Neave, Manning, & Grammer, 2006; Jones, Little, Feinberg, et al., 2004; Rhodes et al., 2001) and correlated positively with childhood health (based on detailed medical records; Rhodes et al., 2001). In summary, facial averageness is a fixed cue to attractiveness, independent of symmetry, and has been shown to be related to both perceived and actual health.

1.3 Sexual Dimorphism

Facial sexual dimorphism, or sex typicality, refers to the masculinity or femininity of a face. Estrogen is largely related to feminine facial features, while androgens such as testosterone govern the development of masculine facial features (Fink & Penton-Voak, 2002; Law Smith et al., 2006; Miller & Todd, 1998). Males and females show significant differences in overall face size, eyebrow ridges, and jawbone size, with women having less prominent eyebrow ridges and a smaller chin compared to men (Enlow & Hans, 1996; Weston, Friday, & Liò, 2007).

Some research has found that women preferred feminized male faces (Little & Hancock, 2002; Perrett et al., 1998), while other researchers have found a preference for masculinized male faces (Cunningham, Barbee, & Pike, 1990; Grammer & Thornhill, 1994; Scheib et al., 1999). Women prefer more masculine faces in a short-term relationship context and more feminine faces in a long-term relationship context (Conway, Jones, DeBruine, & Little, 2010; Little, Cohen, Jones, & Belsky, 2007). Research also suggests women’s preferences for masculinity can change according to their menstrual cycle, with a
higher preference for masculinity in the fertile phase and a higher preference for femininity during the non-fertile phase (Jones et al., 2008; Little & Jones, 2012; Penton-Voak & Perrett, 2000; Penton-Voak et al., 1999). Indeed, differences in culture (Penton-Voak, Jacobson, & Trivers, 2004), pathogen disgust (DeBruine, Jones, Tybur, Lieberman, & Griskevicius, 2010), and access to medical care (DeBruine, Jones, Crawford, Welling, & Little, 2010; Penton-Voak et al., 2004) can also influence preference for masculinity. In turn, these findings suggest female perceptions of masculinity are complex and many variables should be considered when referring to the attractiveness of masculinity.

Within female faces, there is general agreement that femininity (Jones et al., 1995; Perrett et al., 1998) is attractive, which is consistent with findings that female facial femininity correlates with youthfulness (Jones et al., 1995), reproductive hormone levels (Law Smith et al., 2006) and maternal tendencies (Law Smith et al., 2012).

Measurements of masculinity, like other facial cues discussed previously, have been widely disputed in the literature. While some researchers argue that simply asking individuals to rate perceived masculinity is enough to capture differences in sexual dimorphism as perceived by raters (e.g., Rhodes, 2006) others have contended that perceptual ratings of masculinity can be confounded by other parameters (i.e., facial averageness and/or symmetry), as well as sexual stereotypes of sex-specific personality traits. In light of this, many researchers advocate using morphological measures of masculinity (Komori, Kawamura, & Ishihara, 2011). While perceived masculinity is linked to perceived attractiveness (Koehler, Simmons, Rhodes, & Peters, 2004; Rhodes et al., 2007; Rhodes, Simmons, & Peters, 2005), other studies that use morphological measurements of masculinity instead, fail to replicate a relationship between masculinity and attractiveness (Koehler et al., 2004; Penton-Voak et al., 2001; Stephen et al., 2012; Thornhill & Gangestad, 2006; Waynforth, Delwadia, & Camm, 2005). The differences in findings for perceived and measured facial cues are noteworthy, and will be further investigated in the context of eyelid-openness and "neutral" mouth curvature in Chapter 4.

In keeping with the relationship between testosterone and facial masculinity, both perceived and measured aspects of facial masculinity have been shown to correlate with handgrip strength (Fink, Neave, & Seydel, 2007; Windhager, Schaefer, & Fink, 2011) and testosterone (Lefevre, Lewis, Perrett, & Penke, 2013; Penton-Voak & Chen, 2004; Roney,
Hanson, Durante, & Maestripieri, 2006). While some studies have suggested facial masculinity is associated with actual aggression (Carré & McCormick, 2008) and actual dominance (Haselhuhn & Wong, 2012; Lewis, Lefevre, & Bates, 2012; Stirrat & Perrett, 2010), more recent studies with larger samples have failed to replicate findings (Gómez-Valdés et al., 2013; Lefevre et al., 2012; Özener, 2012). Nonetheless, individuals usually perceive masculine faces as aggressive and dominant (Alrajih & Ward, 2013; Carré, McCormick, & Mondloch, 2009; Carré, Morrissey, Mondloch, & McCormick, 2010; Geniole, Keyes, Mondloch, Carré, & McCormick, 2012; Short et al., 2012). Perrett et al. (1998) also found that masculinized faces were rated as less kind, less emotional, colder, less honest, less cooperative, and less likely to be a good parent, while feminized faces were rated in an opposite manner.

The relationship between masculinity and health has been controversial. Numerous studies have found a relationship between masculinity and perceived health (Boothroyd, Scott, Gray, Coombes, & Pound, 2013; Johnston, Hagel, Franklin, Fink, & Grammer, 2001; Rhodes, Chan, Zebrowitz, & Simmons, 2003; Rhodes, 2006; Scott, Swami, Josephson, & Penton-Voak, 2008; Smith, Jones, Debruine, & Little, 2009; Zebrowitz & Rhodes, 2004), yet studies investigating the relationship between masculinity and actual health are inconclusive (Boothroyd et al., 2013; Rhodes et al., 2003; Roberts, Buchanan, & Evans, 2004; Scott, Clark, Boothroyd, & Penton-Voak, 2013). The relationship between masculinity and health has been debated (for discussion see Boothroyd et al., 2013; Rantala et al., 2013; Scott et al., 2013), but these debates go beyond the scope of this thesis. More pertinently, the relationship between facial masculinity and attractiveness is complex given the simultaneous links of masculinity to negative personality attributions, but potentially positive health attributions.

### 1.4 Skin Condition

Several studies have found a relationship between skin condition (skin texture and skin color) and perceived attractiveness. People typically prefer skin that is free from lesions and atypical growths (Symons, 1995). Further, people typically perceive evenness in skin texture (Fink, Grammer, & Thornhill, 2001) and skin color (Fink, Grammer, & Matts, 2006; Matts, Fink, Grammer, & Burquest, 2007) as younger, healthier, and more
attractive. Additionally, research has found that skin patches (on cheeks) from symmetrical faces were perceived as more healthy than skin patches from asymmetric faces (Jones, Little, Burt, & Perrett, 2004). The same study showed that digitally putting skin texture that had previously been rated as healthy onto other faces made those faces appear more attractive than faces with unhealthy rated skin texture superimposed onto them (Jones et al., 2004).

Skin condition is related to health, and the skin attributes that people find attractive correlate with good health. Jones et al. (2004) argue that perceptions of health mediate the relationship between skin color and attractiveness, such that skin color affects facial attractiveness through the influence of color on perceived health.

Research indicates that when participants are asked to manipulate face color to optimize perceived health, participants increase skin redness, yellowness, and lightness (Stephen, Coetzee, et al., 2009; Stephen, Law Smith, Stirrat, & Perrett, 2009). These three dimensions of skin color preference may each reflect different aspects of health. More skin redness indicates oxygenated blood levels that increase with respiratory health (Armstrong & Welsman, 2001) and decrease with cardiac and respiratory illness (Ponsonby, Dwyer, & Couper, 1997). Similarly, more skin yellowness correlates with higher carotenoid consumption (Edwards & Duntley, 1939; Stamatas, Zmudzka, Kollias, & Beer, 2004; von Schantz, Bensch, Grahn, Hasselquist, & Wittzell, 1999). Carotenoids are antioxidants obtained in diet (through fruits and vegetables) and expended in immune defense (Alaluf, Heinrich, Stahl, Tronnier, & Wiseman, 2002; Friis et al., 2001). Hence, high skin yellowness may signal a healthy state and a good diet (Stephen, Coetzee, et al., 2009; Stephen, Law Smith, et al., 2009). Skin lightness is mainly influenced by melanin, with higher melanin making skin darker (and also more yellow). While some melanin is healthy, as it has been shown to prevent skin cancer and sunburn (Robins, 1991), as well as fetal development defects (Omaye, 1993), too much melanin can inhibit vitamin D synthesis and thus have negative effects on health. In summary, there is a growing realization that the malleable facial cue of skin color plays a large role in perceived health and attractiveness, a role that is greater than fixed facial traits such as sexual dimorphism (Scott et al., 2013; Stephen et al., 2012).
1.5 Facial Adiposity

Perceptions of adiposity, the perception of weight in the face (Coetzee et al., 2009), show a connection between perceived attractiveness and health similar to that of facial symmetry, averageness, sexual dimorphism, and skin condition. Yet, like skin condition, but unlike facial symmetry, averageness, and sexual dimorphism, adiposity can change over a fairly short time scale.

The relationship between adiposity, health, and attractiveness is also similar to the relationship between body mass index (BMI), attractiveness of body shape (Parkinson, Tovée, & Cohen-Tovee, 1998; Swami et al., 2007; Swami & Tovée, 2005; Tovée, Reinhardt, Emery, & Cornelissen, 1998), and health (Flegal, Graubard, Williamson, & Gail, 2005; Manson et al., 1995; Mokdad et al., 2003; Must et al., 1999; Pi-Sunyer, 1993; Ritz & Gardner, 2006; Wilson, D’Agostino, Sullivan, Parise, & Kannel, 2002). Specifically, the relationship between BMI and health is curvilinear, with both very high and very low BMI being unhealthy, which matches the curvilinear relationship between BMI and perceived attractiveness of body shape (i.e., very high and very low BMI is unattractive; Furnham & Baguma, 1994; Maisey, Vale, Cornelissen, & Tovée, 1999; Tovée et al., 1998). Indeed, Coetzee et al., (2009) find a similar curvilinear relationship between perceived adiposity and perceived health and attractiveness.

There has been a growing amount of research suggesting a relationship between perceived adiposity and actual health. Research has found that adiposity was a significant predictor of actual health measured by respiratory infections, use of antibiotics, and blood pressure (Coetzee et al., 2009). Additionally, research has found a relationship between perceived weight and a composite measure of physical and psychological health (Tinlin et al., 2013). In men, adiposity has also been linked to an antibody response to the hepatitis B vaccination and was shown to mediate the relationship between antibody response and perceived attractiveness (Rantala et al., 2013).

1.6 Interacting Cues
All in all, there are multiple facial cues to attractiveness and health, some of which change slowly over time depending on maturity (masculinity) or developmental stability (averageness and symmetry), while others change on a quicker time scale according to diet and exercise (skin condition and adiposity). There is increasing evidence that the importance of malleable cues compared to fixed traits in attractiveness has been underestimated.

In general, separate cues to attractiveness covary, such that male facial masculinity and male facial symmetry are positively associated (Little, Jones, et al., 2008), and high facial symmetry is associated with an attractive skin appearance (Jones, Little, Burt, & Perrett, 2004). The positive correlations amongst attractiveness cues provide further evidence for a general "fitness factor" (Candolin, 2003; Feinberg et al., 2005; Thornhill & Grammer, 1999).

The interaction amongst facial cues is also noteworthy. Some have argued that adiposity is a more significant predictor of health (measured by hepatitis B antibody response) than masculinity (Rantala et al., 2013). Fisher, Han, DeBruine, and Jones (2014) found a significant interaction between adiposity and color cues, such that low adiposity was perceived as more attractive and healthy for faces with increased skin redness and yellowness (healthy-looking coloration); whereas adiposity had less of an impact on ratings of unhealthily colored faces. Stephen (2012) found that facial skin color is a better predictor of attractiveness than facial masculinity. These findings are consistent with findings from Stephen et al. (2012) suggesting that "state" cues, which are cues related to current or recent health, are more important predictors of attractiveness than "trait" cues, which are more structural cues to past health during development. In turn, it seems that perceptions of attractiveness and perceived health are complex, in that they reflect and perhaps overgeneralize from cues to current and previous health.

1.7 Stereotypes and Overgeneralizations

The tendency to judge others by their facial appearance and the consensus shown in such judgments is noteworthy and begs the question of the origins of such impression formation. The literature on stereotypes provides some insight on consensus among judgments of facial stimuli. While literature on stereotypes is varied, researchers tend to
share the perspective that stereotypes are related to three underlying principles: stereotypes are aids to explanation, stereotypes are energy-saving devices, and stereotypes are shared group beliefs (Devine, 1989; McGarty, Yzerbyt, & Spears, 2002). Thus, the tendency to use stereotypes when making judgments on facial appearance seems appropriate in "aiding explanation" when no other information about the target is available. Early research on stereotypes by Katz and Braly (1935) defined stereotypes as "fixed impressions which conform very little to the facts it pretends to represent" (p. 181). Yet, more recent definitions of stereotypes have highlighted that stereotypes need not be accurate or negative and that stereotypes are beliefs about the characteristics or attributes of a group (Judd & Park, 1993). For example, one may hold the stereotype that attractive people are more healthy than unattractive people.

Overgeneralizations can be seen as using stereotypes (e.g., attractive people are healthier) to form impressions or guide perceptions. Specifically, overgeneralization can be defined as a logical fallacy that occurs when a conclusion about a group is drawn from an unrepresentative (too small or narrow) sample (Walton, 1999). Thus, overgeneralizations, like stereotypes, need not be accurate, but given the nature of drawing conclusions from unrepresentative samples, they probably are inaccurate. Zebrowitz, Fellous, Mignault, and Andreoletti (2003) argued that overgeneralizations involve accurate trait impressions in certain faces (a small unrepresentative sample) being overgeneralized to other faces that are similar to the unrepresentative sample. Zebrowitz et al. (2003) extend the overgeneralization hypothesis to the anomalous face overgeneralization in which faces resembling those that are "unfit" illicit negative responses. For example, the perception of normal individuals with uneven and pale skin color as unattractive and unhealthy may be an overgeneralized response based on an adaptation to avoid illness or contagion, as some individuals who show marked skin paleness and unevenness are, in fact, unhealthy (Stephen, Coetzee, et al., 2009). Arguably, such overgeneralizations could be seen as maladaptive, given that they lead individuals to reject healthy and fertile, yet unattractive, individuals as mates. On the other hand, ecological theory suggests that the cost associated with failure to avoid such ill health information can be more maladaptive than missing healthy and fertile mates (Zebrowitz et al., 2003).

When discussing overgeneralizations, the issue of accuracy comes into play, for overgeneralizations typically stem from a belief, which by nature, is thought to be accurate
(at least by the perceiver). A thorough review by Todorov et al. (2015) highlights several factors that emphasize the overstated accuracy of trait attributions from faces alone. In particular, research has found that failing to properly control for obvious indicators of the traits being deduced (i.e., target’s gender, age, and ethnicity) may result in an overestimation of how much information is accurately being inferred from the face alone (Olivola, Eubanks, & Lovelace, 2014; Olivola, Sussman, Tsetsos, Kang, & Todorov, 2012). For instance, amongst children and adolescents, those with older looking faces are likely to be perceived as more intelligent, which is congruent with increases in actual crystalized intelligence with age (discussed more in Chapter 2). Another factor that may influence the overestimation of accuracy in attributions from faces is the significant differences in images of the same person taken on different occasions (Todorov & Porter, 2014). For example, a person with a subtle smile may be perceived as more extroverted, but an image of the same person with a subtle frown might result in different perceptions of personality and perhaps different accuracy in perceived personality.

Subtle smiling, sometimes referred to as positive emotional valence, has been shown to have a marked influence on impression formation. From behavioral studies and computer modeling, Oosterholf and Todorov (2008) introduced a two-dimensional model demonstrating an underlying structure for face evaluations based on two primary dimensions: trustworthiness or valence and dominance. That is, while many attributions (intelligence, trustworthiness, extraversion, neuroticism, agreeableness, etc.) can be made to faces, all of these attributions can be accounted for by just differing amounts of the two dimensions. While Todorov and colleagues (2008) worked with faces with supposedly neutral expressions, subtle shape differences in mouth curvature reminiscent of smiling and frowning exerted a profound influence on impression of trustworthiness. Findings suggest that judgments of trustworthiness stem from an overgeneralization made to emotional expressions that may signal hostile or friendly intentions, which can be related to whether the person should be avoided or approached. Specifically, the valence dimension accounted for most of the variance in a variety of trait judgments (63.3%) which were mainly related to positive judgments such as whether a face is perceived as attractive or responsible. On the other hand, the dominance dimension is more sensitive to features signaling physical strength (e.g., masculinity and maturity of age) and related to judgments such as aggressiveness and confidence. Sutherland et al. (2013) replicated the two-
dimensional valence/trustworthiness by dominance model of social inferences and added a new “youthful-attractiveness” factor. Researchers reasoned that the emergence of the new “youthful-attractiveness” dimension was related to differences in a more realistic set of stimuli used, yet it could also be argued that such unstandardized stimuli (e.g., overt expressions, dramatic differences in head posture, etc.) may confuse interpretations.

The dimensions of dominance and trustworthiness/valence are particularly interesting considering their relatedness to fixed (or static) and malleable (or dynamic) cues to trait attributions. Hehman, Flake, and Freeman (2015) discuss the differences in the influence of static and dynamic facial cues on the attributions related to dimensions of trustworthiness/valence and dominance. Specifically, Hehman et al. (2015) argue that dynamic facial cues are more likely to influence judgments of intentions (trustworthiness/valence dimension), while evaluations of ability (dominance/physical ability dimension) are more likely to rely on static facial cues. For example, the degree to which a person is smiling and has eyes wide open can change quickly and in turn influences their perceived mood or tiredness accordingly. Similarly, an overgeneralization effect may result from subtle cues to tiredness influencing perceptions of intelligence, given the well-known relationship between tiredness and reduced cognitive ability (Ohayon & Vecchierini, 2005; Pilcher & Huffcutt, 1996; Wickens, Hutchins, Laux, & Sebok, 2015). Chapter 4 investigates the potential for dynamic cues of mouth curvature and eyelid-openness to influence perceptions of intelligence accordingly.

1.8 Summary

Various studies have investigated facial cues and their influence on attractiveness and other personality trait attributions. It is clear that facial symmetry, averageness, sexual dimorphism, skin color, and adiposity all have at least some biological basis. These facial cues seem to offer a potential basis for some accuracy in attributions to the actual state of a person’s health and fitness, testosterone level, blood perfusion and oxygenation state, carotenoid levels, and BMI, yet people overgeneralize these facial cues to guide their social attributions of personality, strength, aggression, and health. Literature suggests these perceptions all seem to be related to two underlying dimensions: trustworthiness/valence and dominance (Oosterhof & Todorov, 2008), which are of particular importance when
exploring the role of fixed and malleable facial cues on social judgments (Hehman et al., 2015).

The different methods used to measure the facial cues discussed have all improved to some extent over time, based on emerging evidence and technological advancements. Yet, different methods have often resulted in different, and at times, conflicting findings and interpretations; it is important, therefore, to advance methodological approaches to understand the role of certain facial cues on first impressions.

In an attempt to offer a novel measurement of malleable facial cues that may have been underestimated, Chapter 4 explores the role of eyelid-openness and mouth curvature on impressions of attractiveness and intelligence. Specifically, we created new measurements that can detect slight differences in eyelid-openness and mouth curvature in facial images and validated these measurements with perceptual judgments from raters. Through mediation analysis we investigated the role of these malleable cues on perceptions of intelligence, independent of attractiveness. In a similar manner to the other cues discussed thus far, Chapter 4 discusses the potential validity of cues of eyelid-openness and mouth curvature by reviewing their relationship to actual sleep deprivation, and thus exploring the overgeneralizations that may follow.
Chapter 2: What Is Intelligence?
2. Intelligence

Beliefs about intelligence influence an evaluator’s perceptions of what makes a face intelligent looking. This chapter focuses on how varying perspectives and understandings of the effects of intelligence on life outcomes (i.e., mate selection and academic performance) may influence accurate perceptions of competence, while Chapter 3 will go on to discuss how individual differences may influence the extent to which an individual endorses the perceived intelligence-attractiveness relationship.

The Merriam-Webster dictionary defines intelligence as "the ability to learn or understand things or deal with new or difficult situations" (Intelligence, 2016), yet theories and definitions of intelligence differ, and the nature of intelligence is arguably one of the most controversial topics in psychology, in large part due to its debated relationship to socioeconomic status (Capron & Duyme, 1989; Jencks, 1972; Mackintosh, 1998) and ethnicity (Loehlin, 1974; Mackintosh, 1998; Rushton & Jensen, 2005). This thesis mitigates the problem of in-group favoritism (Zebrowitz, Bronstad, & Lee, 2007) by examining perceptions of Caucasian faces by Caucasian evaluators exclusively. Yet, different views on definitions and supposed outcomes of intelligence may still sway the degree to which competence is perceived accurately.

After broadly discussing theories and definitions of intelligence, this chapter will discuss potential consequences of intelligence such as resource acquisition, mate selection, and academic performance. Theories and outcomes of intelligence serve as an overview primarily for Chapter 6, which investigates whether individuals can accurately assess academic performance from faces alone when controlling for the well-documented attractiveness halo (discussed more in Chapter 3) and examines the differences in accuracy amongst different perceived competence measures.

2.1 Theories of Intelligence

An individual’s implicit theory of intelligence refers to an individual’s fundamental underlying beliefs about whether intelligence or abilities can change or are static (Dweck, Chiu, & Hong, 1995; Dweck & Henderson, 1989). Individuals who believe
intelligence and ability are fixed and unchangeable subscribe to the entity mindset (sometimes referred to as a fixed mindset), while an individual who believes hard work and effort influences intelligence has an incremental mindset (sometimes referred to as a growth mindset). These implicit theories influence behavior; individuals who subscribe to incremental theory perform better academically (Blackwell, Trzesniewski, & Dweck, 2007) and have stronger negotiating skills (Kray & Haselhuhn, 2007). Thus far, research on implicit theories of intelligence has focused on impacts on goal orientation and motivation. Yet, there are reasons to believe that implicit theories of intelligence could also impact upon perceptions of other people’s intelligence, for example they may alter the accuracy of judgements of the intelligence of others.

### 2.1.1 Multiple intelligences

Individuals (including psychological researchers), differ in their definitions of intelligence. Some researchers have criticized the traditional conception of intelligence, as measured by standardized intelligence tests, by arguing that intelligence tests primarily measure how “book smart” an individual is. In turn, various researchers have proposed that there are multiple intelligences (Gardner, 1983, 2000; Sternberg, 1985, 2000). Guilford (1959, 1977) pioneered the concept of separate measures of intelligence, suggesting human mental abilities encompass about 150 different types. Later, Gardner (1983, 2000) suggested that there are eight different types of intelligence: rhythmic, spatial, linguistic, mathematical, kinesthetic, interpersonal, intrapersonal, and naturalistic. Thus, while someone may be musically gifted, they may be less mathematically inclined. Theories of multiple intelligences have inherent appeal in that they imply that everyone is smart at something, even if the traditional intelligence test (e.g., Raven Standard Progressive Matrices, Wechsler Adult Intelligence Test, etc.) does not reflect it. Yet, less empirical evidence supports multiple intelligences theories compared to other theories of intelligence (Carroll, 1993; Gottfredson, 2003).

### 2.1.2 General factor g

Psychometrics and traditional psychology appreciates Spearman’s (1905) g factor of general intellectual ability, which is the underlying performance of an individual based on scores on a range of abilities. In essence, Spearman’s g suggests that different ability
tests are significantly intercorrelated; for example, someone who scores high on a verbal test will likely score similarly high on a numerical test.

Spearman (1927) developed a data-reduction procedure called factor analysis that examines the correlations among a series of measures to identify underlying patterns in the data that can be attributed to an unobservable variable (also called a latent variable). Through this method, Spearman (1927) found that performance on various aspects of cognitive ability could be represented by a single general underlying mental ability. Spearman (1927) also developed statistical methods to correct for attenuation as a means to improve measurements of intelligence. This method accounts for measurement error when calculating the correlation between variables. Since cognitive/intelligence tests are not completely accurate in the measure they aim to assess (all tests have a certain level of measurement error), Spearman developed a formula that considers the unreliability of variables being correlated to improve estimates of the relationship between the variables (Spearman, 1927). This method of correcting for attenuation is widely used by psychologists and is used in Chapter 4 and 6 of this thesis.

### 2.1.3 Fluid and crystalized intelligence

One of Spearman’s PhD students, Cattell (1963), extended Spearman’s (1905, 1927) findings to create the theory of crystallized and fluid intelligence. Using similar factor analysis techniques, Cattell (1963) argued that general intelligence can be categorized into fluid intelligence and crystalized intelligence. Fluid intelligence enables an individual to perform well on nonverbal tasks and other measures of culture-free cognitive performance that do not require previous knowledge. Crystalized intelligence enables an individual to do well on verbal tasks and is substantially influenced by learning and previous knowledge. Chamorro-Prezmuic (2007) described the difference between fluid and crystalized intelligence using an analogy with computers, saying that fluid intelligence is similar to the processor or hardware of the computer, while crystalized intelligence is more akin to the software on the computer. Testing Cattell’s (1963) theory, Horn and Cattell (1967) found that younger adults had higher fluid intelligence while older adults had higher crystallized intelligence, suggesting the negative effects of normal aging on fluid intelligence and the positive effects of additional learning and enculturation on crystalized intelligence. Horn and Cattell’s (1967) findings emphasize the importance of
breaking down intelligence by suggesting that, depending on what type of intelligence is being measured (fluid or crystalized), intelligence can either increase or decrease with age (p. 124).

There are various measures that assess fluid intelligence in a way that is aimed to be culturally unbiased (Cattell, Feingold, & Sarason, 1941; Cattell, 1940). Some of the most popular, privately owned tests are the Cattell Culture Fair IQ Test, the Raven’s Progressive Matrices, and the Wechsler Adult Intelligence Test. There are many types of intelligence test items and these have become much more sophisticated over time, with some of the first tests measuring only reaction time to a sound and time for naming colors (Cattell & Galton, 1890). Some of the most common subtests of general intelligence nowadays are: mental rotation tasks which attempt to measure spatial cognition and the rate of processing of spatial information; letter and number series that assess logical reasoning; matrix reasoning items, which assess nonverbal analytical ability; and verbal reasoning, which measures comprehension and critical reasoning (Chamorro-Premuzic, 2007). In the nature of open access science, Condon and Revelle (2014) developed the International Cognitive Ability Resource (ICAR), a 16-item intelligence test and made this resource freely available to researchers interested in measuring intelligence. ICAR is an unproctored and untimed assessment that has been shown to correlate moderately to strongly with other measures of cognitive ability and achievement and is used in Chapter 5. See Appendix 1 for examples on the different types of intelligence test items discussed (i.e., mental rotation, matrix reasoning, letter and number series, and verbal reasoning) and used in Chapter 5.

### 2.1.4 Defining intelligence

Theories on single general factor intelligence and multiple intelligences may both be correct; that is, they are not mutually exclusive. A single underlying intelligence may lead to strong performance on multiple measures, but these measurements may still represent distinct types of cognitive abilities. Nonetheless, varying perspectives, measurements, and theories of intelligence have led to various definitions of intelligence. In turn, a group of researchers ranging in experience and opinions collaborated on a report (one of the most cited of its kind) entitled *Intelligence: Knowns and Unknowns* which
reviewed approaches to intelligence (Neisser et al., 1996). The following is an excerpt from the report:

"Individuals differ from one another in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought. Although these individual differences can be substantial, they are never entirely consistent: a given person's intellectual performance will vary on different occasions, in different domains, as judged by different criteria. Concepts of 'intelligence' are attempts to clarify and organize this complex set of phenomena. Although considerable clarity has been achieved in some areas, no such conceptualization has yet answered all the important questions, and none commands universal assent... Indeed, when two dozen prominent theorists were recently asked to define intelligence, they gave two dozen, somewhat different, definitions" (Neisser et al., 1996, p. 77).

### 2.2 Outcomes of Intelligence

Outcomes of intelligence are noteworthy as they shed light on the importance of intelligence, how intelligence is perceived, and why it might be such a sought after trait. In exploring the potential accuracy of competence from faces alone (Chapter 6), it is of particular interest to investigate the extent to which different forms of competence (e.g., physical competence, diligence, etc.) are attractive and the relationship between different assessments of competence (i.e., academic performance, conscientiousness and intelligence).

#### 2.2.1 Mate selection

Intelligence is highly valued, which is not surprising given the clear implications of cognitive ability in everyday life. Indeed, in an analysis of human mate preferences across 37 cultures, intelligence ranked among the top four desired characteristics in potential mates (Buss, 1989). It seems both men and women favor intelligence equally, as Buss (1989) found that men and women only slightly differed in their preferences for
intelligence in partners in the context of sexual relations (men had lower acceptable intelligence standards in regards to short-term sexual relationships, but both men and women equally preferred intelligence in a partner to date and marry). As discussed in Chapter 1, intelligence is also related to health which, through a general "fitness factor" would also explain its relationship to attractiveness. Women may favor traits like physical strength and assess them through indicators like height or face shape (Krams et al., 2014; Sell et al., 2009). Similarly, individuals may favor intelligence as a trait that benefits the acquisition of resources and assess it through traits like alertness, wit, charisma, and imagination (Prokosch et al., 2009). Additionally, personality traits such as conscientiousness may also be a cue to capacity for resource acquisition, since acquisition may also be assessed through traits like hard work and thoroughness (Chamorro-Premuzic, 2006, 2007). Thus, the attractiveness of perceived competence measures, which may signal resource acquisition ability, overlap.

2.2.2 Academic performance

Given that intelligence tests are validated by their ability to predict academic success, it is no surprise that intelligence predicts academic performance (Chamorro-Premuzic, 2007). As such, an intelligence test that did not correlate with achievement at school probably would not be labeled as an intelligence test. While intelligence is a strong predictor of academic performance, other variables, like personality, have been found to be even stronger predictors of academic performance (Chamorro-Premuzic, 2007). Specifically, conscientiousness, described as thoroughness, carefulness, organized, and efficient, has consistently shown to be a stronger predictor of academic performance than intelligence (Moutafi, Furnham, & Paltiel, 2004; Roberts, Lejuez, Krueger, Richards, & Hill, 2014; Wood & Englert, 2009).

Moreover, the predictive power of intelligence on academic performance seems to decrease as students reach higher academic levels (Furnham & Bachtiar, 2008; Sanders, Osborne, & Greene, 1955). Researchers argued this is probably due to different forms of assessments (continuous assessments and coursework over long periods of time) and restrictions in range of intelligence (students that go to university are smarter as a group than high school students, making intelligence more homogeneous). Consequently, research suggests that conscientiousness is a stronger predictor of academic performance in
higher education than intelligence (Chamorro-Premuzic & Arteche, 2008; Chamorro-Premuzic, 2006). Accordingly, Chapter 6 investigates the overlap between the attractiveness of intelligence and conscientiousness, as well as the predictive power of conscientiousness over intelligence in predicting academic performance.

2.3 Summary

Given the multifaceted debate among experts in the field of intelligence, there are various definitions of intelligence. Some people may think that intelligence is fixed and does not change considerably (like fluid intelligence); while others may think it accumulates over time (like crystalized intelligence). Likewise, some might agree on an incremental or entity theory of intelligence, but hold significantly different opinions on whether there is one general type of intelligence or multiple types of intelligence (or both).

Correspondingly, individuals may differ in the extent to which they believe intelligence predicts academic performance more than hard work or effort. For example, some might argue that it does not matter how much an unintelligent person works; they would be limited in how well they do academically (and indeed this may be true in the extreme lower end of the spectrum marked by learning disabilities). It is possible that each of these ambiguities in the term “intelligence” may limit accuracy or consensus in predicting academic performance from facial appearance. By contrast, other aspects of competence and personality may be less ambiguous.

Chapter 6 assesses the potential for individuals to accurately assess academic performance from facial appearance alone by asking them to rate perceived intelligence, conscientiousness, and academic performance. Given the strong implications of the attractiveness halo (discussed more in Chapter 3), we explore whether statistically controlling for facial attractiveness would result in more accurate perceptions of academic performance.
Chapter 3: Attractiveness-Intelligence Halo
3. Halo

As discussed in Chapter 1, individual differences in culture, perceived self-worth, access to medical care, and hormones are just some variables that can influence participants’ preference for facial symmetry and sexual dimorphism. In a similar way, individual differences may influence the degree to which individuals endorse the attractiveness-intelligence relationship in faces, a relationship under scrutiny because of its likely effect on education and other situations where expectations shape performance. This chapter will broadly discuss the attractiveness halo effect, the expectancy effect, and the intelligence-attractiveness halo. It will shed light on why some individuals might perceive a stronger association between intelligence and attractiveness than others (further investigated in Chapter 5).

3.1 Attractiveness and Intelligence Halo

A substantial amount of literature on the “attractiveness halo effect” shows that people ascribe desirable personality traits to others they find attractive (Dion et al., 1972). One potential explanation for this effect is the “good genes” hypothesis; that is, that physical attractiveness is a valid indicator of the ability to pass on other genes that will increase survival and reproductive success (Fisher, 1915). As Chapter 1 described, symmetry (Thornhill & Gangestad, 2006), averageness (Langlois & Roggman, 1990), sexually dimorphic features (Perrett et al., 1998, 1999), skin pigmentation and texture (Fink et al., 2001; Stephen, Law Smith, et al., 2009), and adiposity (Coetzee et al., 2009; Rantala et al., 2013) all influence attractiveness, and each attribute has a relationship to health (Rhodes, 2006). Given the convincing amount of research supporting an association between health and cognitive function, cues to health may also be cues to intelligence (discussed further in Chapter 2).

Kanazawa and Kovar (2004) examine the relationship between actual intelligence and attractiveness by presenting four controversial (see Denny, 2008; Mitchem et al., 2014; Zebrowitz & Rhodes, 2004) assumptions: 1) men with higher intelligence scores are more likely to attain higher status than men with lower intelligence scores; 2) higher-status men
are more likely to mate with attractive women than lower-status men; 3) intelligence is heritable; and 4) beauty is heritable (p.228-229). Hence, they conclude, because higher-status men have higher intelligence scores and are more likely to mate with attractive women than lower-status men, their offspring are attractive (because of their mothers) and intelligent (because of their fathers); thus, there is a greater likelihood of an attractive person being intelligent.

Kanazawa and Kovar’s (2004) assumptions have not proven resilient. Denny (2008) points out that there are beautification practices available to members of higher socioeconomic strata. These practices might misrepresent an otherwise less attractive female as more attractive and, therefore, lead the female to be more likely to mate with a higher-status male. Further, research suggests that personality is a stronger indicator of academic and work performance than intelligence (Chamorro-Premuzic, 2007), thus, men of high status do not necessarily have high intelligence (discussed further in Chapter 2). These flawed assumptions perhaps explain the lack of empirical studies replicating Kanazawa and Kovar’s (2004) findings that attractive people are more intelligent (Denny, 2008; Mitchem et al., 2014; Zebrowitz, Hall, Murphy, & Rhodes, 2002; Zebrowitz & Rhodes, 2004).

Correspondingly, Zebrowitz and Rhodes (2004) investigated the relationship between facial attractiveness and actual intelligence in the upper and lower halves of the attractiveness distribution. Results indicated that facial attractiveness and actual intelligence were only significantly correlated in the lower half of the attractiveness distribution (i.e., individuals who were less than average attractive were usually less intelligent). Consistent with the “anomalous face overgeneralization hypothesis,” attractiveness was used (spuriously) as a cue to intelligence across the entire attractiveness distribution, but participants’ judgments of intelligence only applied with any accuracy to faces they considered less attractive (Zebrowitz & Rhodes, 2004). These findings are consistent with the “bad genes” hypothesis, which implies that unattractive faces are a signal of poor genetic fitness, and that faces rated as either average or above average for attractiveness serve as equally relevant signals of genetic fitness.

The most recent of studies on actual intelligence and attractiveness analyzed the largest sample to date (n=1354), utilizing a twin dataset; it found no support for a
relationship between actual intelligence and perceived facial attractiveness (Mitchem et al., 2014). Therefore, while the good (or bad) genes theory may help explain why individuals draw the relationship between attractiveness and perceived intelligence there is not sufficient evidence for a relationship between attractiveness and actual intelligence.

3.2 Attractiveness and the Expectancy Effect

The expectancy effect, demonstrated by the classic Pygmalion study (Rosenthal & Jacobson, 1968), suggests that expectations alone are capable of influencing the target's actual performance. The attractiveness halo has also been studied extensively in relation to its influence on expectations of academic performance (Byrne, London, & Reeves, 1968; Landy & Sigall, 1974; Miller, 1970). It is clear that people ascribe positive educational traits such as intelligence and academic potential to students they consider attractive more strongly than to students they consider unattractive (Dusek & Joseph, 1983; Ritts, Patterson, & Tubbs, 1992). These studies show that variables like gender, race, and past performance affect expectations, but are not significant moderators to the attractiveness halo effect.

Given the influence of the expectancy effect on students' future performance (de Boer, Bosker, & van der Werf, 2010; Rosenthal & Jacobson, 1968) and the influence that attractiveness has on expectations (Dusek & Joseph, 1983; Langlois et al., 2000; Ritts et al., 1992), it is not surprising that students rated as more attractive would do better academically. Indeed, various investigations examining both standardized test scores and grades across different age ranges and grade levels have shown that students rated as more attractive usually receive higher grades and scores on standardized achievement tests than students rated as less attractive (Felson, 1980; Salvia, Algozine, & Sheare, 1977; Singer, 1964). Granted, some empirical studies dispute these findings (Clifford, 1975; Sparacino & Hansell, 1979). The difference in study methods, namely how well a teacher knows the student, may explain inconsistencies in findings; while attractiveness may shape expectations, given sufficient time, other information (e.g., perceived effort, previous academic performance, etc.) may be deterring the influence of attractiveness on perceived intelligence, and in turn, mitigating expectancy effects. For instance, Sparacino and Hansell's (1979) sample of students consisted of a smaller and more intimate class than
other studies have investigated; as they point out, these students may have had the opportunity to influence teacher expectations (and potentially their perceived attractiveness), which would sway the power of the expectancy effect's influence on actual academic performance. Their results indicate the importance of expectancy induction; that is, the process by which teachers raise their expectations of students (Raudenbush, 1984). While expectancy induction could occur through experimental manipulation (i.e., telling teachers which students are smarter), it could also happen naturally by teachers expecting more based on first impressions or other perceived qualities related to academic performance (e.g., perceived intelligence, health, etc.).

Given the consistent and sustained impact of these expectations, the focus of concern has been the accuracy of the expectations themselves. Social psychologists typically argue that teachers’ perceptions of students are usually inaccurate and that teachers’ expectancies influence students’ academic performance to a greater degree than students’ performance influences teachers’ expectancies (Miller & Turnbull, 1986, p. 236). Alvidrez and Weinstein (1999) provide an example of potentially inaccurate expectations influencing academic performance by finding that teacher expectancies formed of children as young as four years old (likely inaccurate expectations given the limited information available at such age) significantly predicted high school performance 14 years later, more so than students’ measured intelligence.

In contrast, educational psychologists have argued that teachers’ expectancies can be accurate (Hoge & Coladarci, 1989), and that inaccurate impressions are typically corrected when more dependable performance information becomes available (Brophy, 1985; Jussim, 1986, 1989). Essentially, the influence of the expectancy effect works through a self-fulfilling prophecy. If, however, the prophecy was based on accurate information, the outcome is likely to reflect correctly perceived ability instead of inaccurate expectations. Jussim and Harber (2005) put it best by pointing out that as accuracy increases, the potential for self-fulfilling prophecies declines; as accuracy decreases, the potential for self-fulfilling prophecies increases (p. 138).

Anderson and Rosenthal (1968) argued that timing affects the power of expectancy induction and that test scores produced before teachers had time to get to know their students had a stronger impact on expectancy effects. Claiborn (1969) concurred, finding
that expectancy induction had no effect on teachers who had already worked with students for about a month. Taken together, these studies imply that while perceptions of attractiveness play a critical role in influencing expectations, the attractiveness bias has more influence in instances of limited context and contact. Furthermore, when educators have access to other information to guide impression formation (e.g., previous academic performance, perceived motivation, etc.) positive educational attributes may act as a halo, influencing perceived attractiveness; this would also lead to a correlation between perceived attractiveness and actual academic performance.

### 3.3 Intelligence and Attractiveness Halo

The general halo effect suggests that perceived intelligence may confer as strong a halo as attractiveness. Yet, the bulk of research has investigated the influence of perceptions of physical attractiveness on impressions of personality and intelligence (de Boer et al., 2010; Dion et al., 1972; Dusek & Joseph, 1983; Langlois et al., 2000; Ritts et al., 1992), and only a few investigate the reverse effect.

Cognitive dissonance theory is one potential explanation for the correlation between perceived intelligence and attractiveness. This theory argues that a high correlation between perceived attractiveness and perceived intelligence may stem from a desire to remain consistent with our initial impressions (Aronson, 1969; Asch, 1946; Festinger, 1962). That is, people will experience increased cognitive consistency if they ascribe attractiveness to an individual to whom they have ascribed other positive traits (Asch, 1946). Dion and colleagues (1972) investigated the attractiveness halo in a paper titled "What is Beautiful is Good," and Gross and Crofton (1977) responded with a paper titled "What is Good is Beautiful," finding that observers perceived female students as more attractive if the observers had received a favorable description of the student’s personality as well as the student’s photograph. Owens and Ford (1978) tested these findings with images of males and females and found similar effects for female but not male images. Two separate studies have argued that external validity for the previous research supporting the attractiveness stereotype is weak (Eagly, Ashmore, Makhijani, & Longo, 1991; Felson & Bohnstedt, 1979). Felson and Bohnstedt (1979) suggest that in a more natural setting, such as one with a more realistic amount of information available,
people are less likely to rely upon appearance in personality or ability judgments, and instead are more likely to use ability judgments to form impressions of attractiveness (see comment by Campbell, 1979 for a more detailed review of the causal relationships deduced and reconciliation with previous experimental results). Thus, the direction in which the halo influences judgments (perceived intelligence influences perceived attractiveness vs. perceived attractiveness influences perceived intelligence) seems dependent on the amount of information available to the perceiver. That is, in situations in which little information is available and perceptions of physical appearance are readily available, perceptions of attractiveness will dominate impression formation until more information is gained.

Kniffin and Wilson (2004) conducted three naturalistic studies and provided evidence for the importance of the differences in impression formation after familiarity, as well as non-physical factors influencing perceptions of physical attractiveness. These researchers found that non-physical factors such as ratings of familiarity, respect, liking, and intelligence, had a significant influence on the perception of physical attractiveness (Kniffin & Wilson, 2004). More recent research by Zhang, Zong, Zhong, and Kou (2014) had participants rate the attractiveness of facial stimuli without any information about their personalities, then rate the same faces two weeks later with positive information, no information, or negative information paired with the facial stimuli. Results indicated support for a "what is good is beautiful" effect in which positive personality influenced increased positive ratings of attractiveness during the second round of rating the same stimuli. Similarly, research has found that group membership can taint perceptions of attractiveness, with followers of a particular group rating their own group leaders as more physically attractive than leaders of competing groups (Kniffin, Wansink, Griskevicius, & Wilson, 2014).

Byrne (1961) indicates a strong correlation between interpersonal attraction and attitude similarity by showing that a participant who believes they share similar attitudes with a stranger will judge the self-similar stranger more positively. Participants will also judge the stranger as better adjusted, better informed, more moral, and more intelligent as compared to a stranger who holds dissimilar attitudes (Byrne, 1961). Park and Schaller (2005) explain that individuals are more likely to express willingness to help a hypothetical person who maintains similar attitudes. Hence, it is possible that perceived similarity in
intelligence would result in a similar effect of biasing future judgments of that similar-intelligence stranger in a positive light.
3.4 Individual Differences

In light of the possibility of an "intelligence halo" in which perceptions of intelligence can influence perceptions of attractiveness, individual differences may influence the degree to which people find intelligence attractive. Indeed, people appear to consider others who resemble themselves as more intelligent or attractive, whether the other person is a stranger (Byrne, 1961) or a potential mate (Buss, 1985; Jensen, 1978; Reynolds, Baker, & Pedersen, 2000).

3.4.1 Assortative mating

As discussed in Chapter 1, the good genes hypothesis attempts to explain how valid signals of actual health (e.g., symmetry, averageness, etc.) can influence perceptions of attractiveness. Given the relationship between actual intelligence and actual health (Bishop, 2009; Eysenck, Derakshan, Santos, & Calvo, 2007; Gregg, 2000; Ohayon & Vecchierini, 2005; Okereke & Grodstein, 2013; Yassuda et al., 2012), it is plausible that honest signals of health may extend to perceptions of intelligence. Assortative mating, defined as the pairing of individuals based on their similarity to each other, may provide insight on individual differences for the degree of attractiveness of signs of "good genes" or indicators of health and intelligence. Humans have been known to assort on inherited attributes such as height and weight (Ginsburg, Livshits, Yakovenko, & Kobyliansky, 1998), as well as non-inherited attributes such as altruism and education (Tognetti, Berticat, Raymond, & Faurie, 2014). Indeed, research has highlighted evidence of assortative mating in a variety of other variables, including age, religion, socioeconomic status, ethnicity, personality traits, and social attitudes (Buss, 1985). Research by Rushton and Nicholson (1988) found evidence suggesting that spouses select each other on the basis of genetic similarity both within and across cultures. Research has found that assortment continues even within groups already selected on the basis of proximity (Rushton, Littlefield, & Lumsden, 1986), indicating that assortment is not simply related to a shared environment.

Based on the literature supporting the assortment of intelligence (Jensen, 1978; Price & Vandenberg, 1980; Reynolds et al., 2000; Vandenberg, 1972) and research on the heritability of general intelligence (Bartels, Rietveld, Van Baal, & Boomsma, 2002; Deary,
Spinath, & Bates, 2006; Neisser et al., 1996; Plomin & Spinath, 2004), Miller (2000) argued that intelligence is sexually, rather than naturally, selected. Hence, an individual’s attraction to a face perceived to be intelligent may partially depend on an observer’s own intelligence.

### 3.4.2 Mate value and standards

In line with the idea that an individual’s intelligence influences perceptions of others’ attractiveness, evidence from several species suggests that individuals with high-quality genes show stronger preferences for good gene markers (i.e., symmetry) than those with lower quality genes (Bakker, Zunzler, & Mazzi, 1999; López, 1999; Mazzi, Künzler, & Bakker, 2003). Women who consider themselves more attractive preferred more masculine faces (a trait that has been argued to signal stronger immunity; see Rhodes, Chan, Zebrowitz, & Simmons, 2003, but also see Scott, Clark, Boothroyd, & Penton-Voak, 2013) and also more symmetric male faces than women who considered themselves less attractive (Kandrik & DeBruine, 2012; Little, Burt, Penton-Voak, & Perrett, 2001; Vukovic et al., 2010). Likewise, females who reported lower self-esteem rated less masculine men as more attractive (Johnston et al., 2001). Accordingly, an evolutionary perspective on perceived mate value and selection standards would suggest individuals with low mate value may adopt lower standards in mate choice, which may reflect an adaptation to avoid the costs of decreased paternal investment or potential abandonment from higher-quality partners (Little et al., 2001; Pawłowski & Dunbar, 1999).

A social psychological perspective would suggest that the importance of one’s self-evaluation in favorability to others may be related to a social exchange or equity model of mate selection (Blau, 1968; Hatfield, Traupmann, & Sprecher, 1985; Hatfield, Walster, Walster, & Berscheid, 1978; Murstein, 1970). These social exchange type models suggest one’s own market-value impacts on one’s trading value, and as such potential partners will attempt to get the best bargain in exchange for their own social assets (Cameron, Oskamp, & Sparks, 1977; Kenrick, 1994; Regan, 1998). Murstein (1970) argued that while seeking a less desirable partner has a low cost (because of a lower risk of rejection), it also has low profit; whereas, seeking a partner who is more desirable than oneself bears high profit and high cost. Thus, the mating game is one of risks, and an accurate self-evaluation
would be advantageous. Indeed, there is evidence to suggest a weak, but significant
correlation between self-rated intelligence and actual intelligence (Furnham & Rawles,
1999; Furnham & Chamorro-Premuzic, 2004; Paulhus, 1998), supporting the possibility
that more intelligent people know their worth and that their preferences for intelligence
will reflect this.

### 3.4.3 Kin selection

Further evidence of a self-similarity preference is kin selection, in which
individuals will favor others that are genetically similar (Hamilton, 1964). Rushton (1988)
extends the concept of kin selection to Genetic Similarity Theory, by suggesting we are not
only biased toward kin, but toward strangers whom we identify as being genetically
similar. In turn, it seems there are various mechanisms that suggest individuals may vary in
the extent to which they find particular traits (like intelligence) attractive and these
individual differences seem related to self-similarity. Theories of kin selection and their
relation to individual differences in perceptions of intelligence will be discussed alongside
Chapter 5, which investigates the role of one’s own intelligence on perceptions of
intelligence in children’s faces.

### 3.5 Summary

For decades, research has explored the influence of the attractiveness halo and its
relationship to perceptions of intelligence and expectancy effects. While some have argued
attractiveness and actual intelligence are linked (Kanazawa, 2004; Kanazawa, 2011;
Zebrowitz et al., 2002), evidence is debated and at best weak (Mitchem et al., 2014).
Nonetheless, literature on the topic has consistently shown evidence to suggest a strong
link between attractiveness and perceived intelligence (Byrne et al., 1968; Dusek &

The influence of perceptions of attractiveness on impressions of intelligence is
particularly worrying, given research showing that teachers’ perceptions of students’
intelligence affects student performance (Anderson & Rosenthal, 1968; de Boer et al.,
2010). While research suggests teachers expect more from students rated as attractive
(Dusek & Joseph, 1983; Ritts et al., 1992), it has also been argued that knowing a student
and being familiar with previous performance records can influence perceptions of

While most research has investigated the effects of attractiveness on impressions of personality and intelligence, less research has investigated the opposite influence, the influence of perceived intelligence on attractiveness. Research suggests non-physical attributes, including intelligence, can influence perceptions of attractiveness (Kniffin & Wilson, 2004). As discussed in Chapter 1, individuals can vary in their preferences for certain physical attributes (i.e., people who prefer extroversion in a mate are more likely to rate extraverted faces as more attractive than introverted faces). Likewise, research on assortative mating, kin selection, mate value, and genetic similarity theories suggest individuals often favor others who are self-similar, which suggests individuals may differ in the degree to which they find certain non-physical attributes (like intelligence) attractive. Given evidence to suggest a self-similarity bias, Chapter 5 investigates whether an individual's own intelligence influences the extent to which an individual rates a face he or she perceives to be intelligent as attractive.
Overview of Experimental Chapters

As discussed in Chapter 1, the face provides cues to various attributes such as health, aggression, and strength (to name a few). Yet, individuals often overgeneralize from facial cues to form impressions. Given the relationship between eyelid-openness and mouth curvature in relation to alertness and mood, Chapter 4 investigates the potential for these cues to influence perceptions of intelligence.

Chapter 3 discusses the strong relationship between perceived attractiveness and intelligence, which emphasizes the need to investigate whether the cues of eyelid-openness and mouth curvature effect perceptions of intelligence directly or indirectly through an influence on perceptions of attractiveness. Using mediation analysis, we investigated the impact of eyelid-openness and mouth curvature across adult (Study 1) and children stimuli (Study 2), as well as through digital manipulation of these cues alone (Study 3) and natural variation in images of individuals before and after sleep restriction (Study 4).

While Chapter 4 investigates the relationship between perceived intelligence and attractiveness at the stimuli level, Chapter 5 explores which individuals endorse the perceived intelligence-attractiveness relationship more than others. Given the self-similarity biases discussed in Chapter 3 (and discussed more in Chapter 5), it was predicted that individuals with higher intelligence would show a stronger endorsement of the perceived intelligence-attractiveness relationship. While Chapter 4 investigates the relationship between facial cues and the attractiveness and intelligence halo and Chapter 5 investigates the role of one’s own intelligence on the degree of endorsement of the perceived intelligence-attractiveness halo, Chapter 6 will assess whether individuals would be more accurate in perceiving academic performance from faces when statistically controlling for attractiveness bias and being more specific about perceived competence tasks.
Chapter 4: Perceived Intelligence: Beyond the Attractiveness Halo

This chapter is based on research that is under peer-review in an academic journal:
4. Abstract

Impression formation is profoundly influenced by facial attractiveness, but the existence of facial cues that affect judgments beyond such an “attractiveness halo” may be underestimated. Over four studies, we investigated malleable facial cues that may influence first impressions of intelligence. In Studies 1 and 2, we scrutinize the perceived intelligence and attractiveness ratings of images of 100 adults (aged 18 - 33) and 90 school-aged children (aged 5 - 17) respectively. Intelligence impression was partially mediated by attractiveness, but independent effects of eyelid-openness and subtle smiling were found that enhanced intelligence ratings independent of attractiveness. To investigate whether these cues were specifically and independently influencing perceptions of intelligence, in Study 3 we digitally manipulated stimuli to have exaggerated eyelid-openness and mouth curvature and found that each independent manipulation had an influence on perceptions of intelligence. In a final set of stimuli (Study 4), we explored changes in these cues before and after sleep restriction, to examine whether natural variations in these cues according to sleep condition can influence perceptions. In Studies 3 and 4 variations in eyelid-openness and mouth curvature were found to influence intelligence ratings. These findings suggest potential overgeneralizations based on subtle facial cues that indicate mood and tiredness, both of which alter cognitive ability. These findings have important implications for students who are directly influenced by expectations of ability and teachers who may form expectations based on initial perceptions of intelligence.
4.1 Introduction

As discussed in the introduction (see Chapter 3), physical attractiveness has long been claimed to produce a favorable bias. In many studies of the attractiveness halo, attractiveness is assumed to be a stable characteristic of faces, but we argue for a focus on more dynamic and malleable facial cues. Malleable cues can also operate as a form of halo bias, but are amenable to change, thus allowing some control over the halo effect.

Considering first existing work on the attractiveness halo, perceptions of attractiveness are strongly correlated to perceptions of intelligence (Eagly et al., 1991; Langlois et al., 2000; Ritts et al., 1992). Rosenthal and Jacobson's (1968) Pygmalion study showed that teacher expectations of intelligence affect students' achievement (see also Babad, Inbar, & Rosenthal, 1982). If an attractive student were perceived as more intelligent, the expectancy effect suggests higher expectations would be placed on attractive students over unattractive students and, through greater attention and other forms of self-fulfilling prophecies, these expectations could in turn influence academic achievement (Clifford & Walster, 1973). Although there are specific exceptions to the rule (such as the "dumb blonde" or less attractive "nerd" stereotypes), there is evidence that the general rule holds more broadly (Ritts et al., 1992).

Specifically, a meta-analytic review by Ritts et al. (1992) found that attractive students are more likely to be ascribed positive educational traits than unattractive students. Attractive students were judged as more intelligent, having more academic potential, better grades, and various other positive educational traits. It was further noted that other variables like gender, race, and past performance also affected expectations, but did not significantly moderate the attractiveness effect. Effects of attractiveness also seem to be sustained over time. Boer, Bosker, and Werf (2010) conducted a longitudinal study consisting of approximately 11,000 students over five years and found that there was a significant relationship between teacher expectation bias and student performance after five years, such that higher expectation by teachers led to sustained higher performance.

Thus, it is clear that, even if later corrected, initial expectations can have important consequences and ratings of facial attractiveness are a known source of initial impression formation (Bar, Neta, & Linz, 2006; Todorov, Mandisodza, Goren, & Hall, 2005;
Zebrowitz, 2004). Langlois, Ritter, Casey, and Sawin (1995) found that even maternal interactions (where familiarity is clearly high) are more positive with attractive children compared to unattractive children. Attractiveness judgments themselves are surprisingly consistent across observers: a meta-analytic review by Langlois et al. (2000) found high inter-rater reliability regarding opinions of facial attractiveness for both men and women, as well as across different countries.

4.1 Static and dynamic facial cues

Research by Hehman, Flake, and Freeman (2015) has highlighted the vital differences between static and dynamic facial cues and their varying impact on assessments of intentions (e.g., trustworthiness, warmth, valence) vs. ability (e.g., physical ability, power). Specifically, it was found that judgments of intentions tend to be based on dynamic facial cues that can change in a short time scale due to the action of facial musculature. Evaluations of ability were found to be more consistent and rely heavily on static facial cues that are dependent on underlying bone structure and hence are less likely to change across multiple instances (Hehman et al., 2015).

Much of the research investigating attractiveness and the “attractiveness halo” has focused on the fixed or static structural facial traits of symmetry, averageness, and sexual dimorphism (Langlois & Roggman, 1990; Perrett et al., 1998, 1999). A “good genes” theory of mate selection argues that these fixed traits are attractive because of their relationship to health (Rhodes, 2006). For instance, a symmetric face may be perceived as attractive because it may denote disease resistance (Thornhill & Gangestad, 2006). An average face may be perceived as attractive because at the psychological level it is less unusual, different, or weird. At the biological level, genetic and environmental stress can produce deviations from averageness and symmetry of the human face (Zebrowitz & Rhodes, 2004, p. 169). Similarly, the immunocompetence-handicap hypothesis (Rhodes et al., 2003) posits that the increased testosterone levels men need to develop sexually dimorphic traits also stress the immune system. Thus, only men with a strong immune system would be capable of withstanding this challenge and are able to grow and sustain masculine traits while remaining healthy. However, the immunocompetence-handicap hypothesis is controversial and recent research supports masculinity only having a limited influence on attractiveness of male faces (Rantala et al., 2013; Scott et al., 2013), and the
effect of masculinity on facial attractiveness is not necessarily consistent across cultures (Stephen et al., 2012). While the described stable traits (symmetry, averageness, masculinity) may all contribute to attractiveness (see Chapter 1), less research has been done on malleable cues to attractiveness and perceived intelligence.

Perhaps students with asymmetric or unique faces may not be doomed to low expectations based on their low perceived attractiveness and hence lowered perceived intelligence. A number of more malleable characteristics have also been shown to affect impressions of attractiveness. For example, skin color and texture are facial cues that can change over relatively short timescales. These skin cues have been found to influence attractiveness and signify current state of health in many species including humans (Fink et al., 2001; Stephen, Law Smith, et al., 2009). We argue here that subtle mouth curvature and eye size cues might also have important impact on perceived attractiveness and intelligence.

### 4.1.1.1 Mouth curvature

Smiling is a malleable facial cue that enhances attractiveness (Conway et al., 2008) and increases the rewarding nature of facial attractiveness (O'Doherty et al., 2003). In turn, smiling has been shown to have a "halo effect" of its own, with the more frequent ascription of desired traits to people who smile over those who do not (Lau, 1982). While a large grin could be interpreted as a naive gesture, we argue that a slight upturn in mouth curvature (i.e., a subtle smile) is likely to influence perceived intelligence in a number of ways. Subtle differences in individual anatomy and facial feature configuration could be reminiscent of emotional expressions. Such hints of expression in supposedly neutral faces are known to affect the attribution of personality traits (Borkenau & Liebler, 1992; Kenny, Horner, Kashy, & Chu, 1992; Kleisner, Chvatalova, & Flegr, 2014; Said, Baron, & Todorov, 2009; Sutherland et al., 2013; Zebrowsitz, Kikuchi, & Fellous, 2010).

Additionally, a sad mood has been shown to significantly reduce cognitive performance (Ellis, Thomas, & Rodriguez, 1984). Someone who is smiling is also likely to be perceived

---

1 The term *mouth curvature* is used throughout most of the paper instead of smiling or frowning because the faces observed were considered neutral; not overtly smiling or frowning. Thus, the subtle differences influencing perceptions are not based on an overt emotion, but rather the shape of the mouth.
as more trustworthy (Oosterhof & Todorov, 2008), approachable, engaged and willing to interact (Jones, Debruine, Little, Conway, & Feinberg, 2006). Consequently, subtle smiling may be a malleable facial cue that enhances perceived intelligence independent of the attractiveness halo.

4.1.1.2 Eye size effects

Subtle smiling and eyelid-openness are characteristics that we will examine in depth as potentially important malleable cues to both attractiveness and intelligence. In exploring eyelid-openness effects, however, we need to consider eye size more broadly, and in particular the relationship between eye size and the phenomenon of neoteny (the attractiveness of more baby-faced features; Guthrie, 1970). Cunningham’s (1986) multiple motives hypothesis of physical attractiveness suggests people are attracted to individuals who display a combination of youthful (neotenuous), mature, and expressive facial features. One fixed feature of neoteny that is considered attractive in both females and males is large eyes (Cunningham, Barbee, & Pike, 1990; Cunningham, 1986).

For our purposes it is critical to note that the measurements used in Cunningham’s (1986) study do not differentiate between eye size and eyelid-openness, whereas that distinction is crucial to our hypotheses. Cunningham’s (1986) measure of eye size is the vertical separation of eyelids relative to vertical height of the face (forehead to chin) multiplied by the average width of each eye relative to the width of the face. We propose a measure of eyelid-openness as a different cue to social judgments as well as a contribution to Cunningham’s (1986) measure of eye size. Indeed, we argue that eye size and eyelid-openness may actually elicit distinct social impressions. Neotenuous features such as large eyes and foreheads, and small chins and noses, may trigger a caring response and the perception of childlike traits including naivety as part of an infant schema (Zebrowitz & Montepare, 1992), whereas eyelid-openness may convey alertness. Taken together, these arguments suggest there can be a conflict between perceived intelligence and perceived attractiveness based on these fixed structural cues to age: people with large eyes should be perceived as attractive, youthful, but potentially naïve and hence less intelligent. Since we are interested in the effects of eyelid-openness over and above the effects of fixed facial features, we adopt a new approach to measuring eyelid-openness that controls for any underlying effects of eye size.
4.1.1.3 Tiredness

Tiredness can also change over a relatively short time scale and has been shown to significantly reduce cognitive performance (Lim & Dinges, 2010; Pilcher & Huffcutt, 1996; Thomas et al., 2000; Walker, 2009). Facial cues of sleep deprivation have been shown to affect perceptions of attractiveness negatively (Axelsson et al., 2010). Sleep-deprived individuals also look sadder than non-sleep-deprived individuals, with corners of the mouth pointing downward (Sundelin et al., 2013). An upturned mouth curvature may thus reflect alertness, but tiredness may also be reflected in the more direct facial cue of eyelid-openness; those who are sleep deprived are likely to have less eyelid-openness than those that are well rested (Sundelin et al., 2013). Additionally, overgeneralization of the relationship between tiredness and impaired cognitive performance may mean that facial cues related to tiredness influence perceptions of intelligence beyond the effects of attractiveness.

4.2 Overview of the Studies

Our overall aim was to examine how malleable facial cues relate to perceived intelligence by focusing on the role of facial cues to mood and alertness. Since both low mood and tiredness can impair cognitive performance, we hypothesized that (malleable) cues to mood and tiredness might affect attributions of intelligence through overgeneralizations. Cues to low mood and low alertness may also have a detrimental impact on attractiveness and therefore such cues could influence attributions of intelligence indirectly through the reduction in attractiveness. Alternatively, the same facial cues may affect impressions of intelligence directly and independently of attractiveness judgments. We were therefore interested in exploring the direct effects of malleable facial cues on intellectual attributions, as well as any indirect effects mediated through the impact on attractiveness.

We began by examining attributions made to adult faces in images chosen to depict a neutral state without any posed expression of emotion. The images employed were from a human photo reference agency and had no overt facial expressions. Since we were particularly interested in investigating the effects of malleable cues in forming teacher expectation during early education, we also replicated the experiment with images of
children's faces with neutral expressions. We then digitally manipulated the mouth curvature and eyelid-openness of stimuli independently to control for any other facial cues or combination of facial cues that could be driving impressions.

Finally, to investigate the malleable nature of these facial cues and to support ecological validity of the digital manipulation findings, we ran a fourth study to determine how subtle changes in facial appearance might alter trait attributions in sleep restricted individuals. We compared pairs of images of the same individuals rested and after sleep restriction. We assumed that, on average, sleep restriction will increase the tiredness of participants and will affect mood adversely. We note that tiredness is a common phenomenon that most adults and children experience and may well affect some individuals often and adversely during their education or employment (e.g., Oginska & Pokorski, 2006; Perkinson-Gloor, Lemola, & Grob, 2013; Perkinson-gloor et al., 2015). Hence, it is likely that cues to tiredness, such as eyelid-openness and subtle frowning will affect attributions and quite possibly expectations of performance.
4.3 Study 1: Adult Faces

To examine cues of perceived intelligence across a stimuli set of 100 facial images, participants were asked to rate facial images of adults with neutral expressions for attractiveness and perceived intelligence. Objective measurements of eyelid-openness and mouth curvature were calculated and validated to assess the influence of these cues to alertness and mood on perception of attractiveness and intelligence. We hypothesized that eyelid-openness and hints of positive affect (upturned mouth curvature) would enhance attribution of both attractiveness and perceived intelligence. We analyzed whether the impact of these facial cues on perceived intelligence was mediated by attractiveness, or whether effects were independent of attractiveness. We also controlled for the fixed trait of eye size, and we expected that large eyes would enhance ratings of attractiveness (Cunningham, 1986), but not necessarily intelligence. Further, we examined the effect of gender since this variable might affect perceived intelligence through biased expectations of male intelligence (Beloff, 1992; Byrd & Stacey, 1993; Furnham, Reeves, & Budhani, 2002) or female attractiveness (Cunningham, 1986; Perrett et al., 1998).

4.3.1 Method

4.3.1.1 Face stimuli

The stimuli were 100 Caucasian two-dimensional faces that included 50 males ($M_{age} = 25.3, SD = 4.64$) and 50 females ($M_{age} = 23.2, SD = 3.74$). These were chosen as the most standardized (e.g., clean shaven, neutral head posture) faces from a commercial database (available at www.3d.sk) designed for use in media and gaming development. An equal number of participants in each gender were purposefully chosen, but age of stimuli was not a part of selection criteria (no particular age range used for exclusions). All photographs chosen for this study were taken with a standardized camera set-up and lighting conditions; individuals had their hair pulled back, did not wear make-up or jewelry, and were instructed to pose with a neutral facial expression.

Using the Psychomorph software, all face images were manually delineated with 188 points (Tiddeman et al., 2001). Points were placed on face feature landmarks (e.g., center of the pupil, eye corners and at regular intervals along the contours of facial
features). Custom digital software was used to then measure the distance between
delineation points for the objective measures calculated. Face images were aligned in size
and position based on left and right pupils. Images were then resized and cropped (1608 x
2584 pixels) so that an equal proportion of hair and neck was shown in each image.

### 4.3.1.2 Objective measurements of facial stimuli

The degree of eyelid-openness was examined by taking the vertical distance from
the center of the pupil to the top eyelid and dividing it by the width of the eye inner
canthus to outer canthus (Figure 4.1). This measure was used because as an eyelid closes
with tiredness, it is the top eyelid that lowers, rather than both eyelids meeting in the
center. Cunningham (1986) measurements of eye size (average width of each eye
relative to the width of face multiplied by the vertical separation of eyelids relative to
vertical height of the face forehead to chin) were also calculated.

![Eyelid-openness calculation](image)

**Figure 4.1.** Eyelid-openness calculation. Eyelid-openness was calculated by taking the
distance from the center of the pupil to the top eyelid (vertical white line) and dividing
it by the width of the eyelid from corner to corner (diagonal white line).

---

2 Slow motion video of eye opening and closing can be seen at [https://youtu.be/okvaOOTvCBw](https://youtu.be/okvaOOTvCBw)
Lastly, despite images supposedly showing neutral facial expressions, measurements of mouth curvature were calculated by taking the average height of the right and left corners of the mouth, subtracting the height of the center of the mouth and dividing by the width of the mouth (Figure 4.2).

Figure 4.2. Mouth curvature measure. Example delineation points marked as A (height of left side corner of mouth), B (height of right corner of mouth), and C (height of center of mouth). The height of the center of the mouth (C) was subtracted from the average height of the corners of the mouth (A & B). This number was then divided by the width of the mouth (distance from A to B). The following formula was used to measure mouth curvature: Mouth corner height relative to center of mouth = \((yA+yB)/2 - yC\); Width of mouth = \(\sqrt{(xA-xB)^2+(yA-yB)^2}\); Mouth curvature = Mouth corner height relative to center of mouth / Width of mouth.

4.3.1.3 Perceptual rating validation of objective measures

Perceptual ratings of smiling and eyelid-openness were collected to validate the physical facial measurements and determine observers' awareness of cues that could potentially influence the attribution of intelligence. With perceptual validation, future
research could use measurements of eyelid-openness and mouth curvature in lieu of perceiver ratings.

A group of 19 own-ethnicity participants (10 female, 9 male, $M_{age} = 27$, $SD = 12.95$) were recruited in an online experiment which (along with the other online studies presented in this thesis) was advertised via the Perception Lab website (www.perceptionlab.com) and Facebook page (www.facebook.com/PerceptionLabStAndrews). No course credit or compensation was allotted. Participants were asked to rate the stimulus faces for degree of eyelid-openness. Given previous work (Elfenbein & Ambady, 2002, 2003) that has highlighted differences in ability to recognize emotions on own- and other-race faces, exclusions were made on the basis of participant ethnicity (four participants excluded; this exclusion criteria was also used in Chapters 5 and 6). Ratings in this task were of a very specific facial feature, hence fewer participants were studied compared to later tasks with more global ratings of attractiveness and perceived intelligence.

Participants first previewed the images, with each face being displayed for one second. After becoming familiar with the stimuli, participants were presented with a scale from 1–7 and received these instructions: "focusing on the eyelids, please rate the degree of eyelid-openness in this face compared to other faces presented." Stimulus order was randomized. Ratings of eyelid-openness for each face were averaged across all participants and correlated with the objective face measurement of eyelid-openness. As expected, the perceptual ratings of eyelid-openness correlated strongly with the physical measurements ($r(100) = 0.78$, $p < 0.001$).

The measure of mouth curvature takes into account both positive and negative reflections of a subtle smile/frown. Therefore in a separate experiment, instead of getting evaluators to rate the degree of smiling on a scale from 1-7$^3$, we emphasized to evaluators the subtlety of expression and allowed evaluators to give a response that reflects either a

---

$^3$ A separate experiment ($n = 23$), in which raters were asked to rate the degree of smiling for the same faces on a scale from 1-7, these perceptual ratings did not correlate significantly with objective measures of mouth curvature ($r(100) = -0.01$, $p = 0.959$). In this experiment, the subtlety of differences in mouth curvature was not emphasized, highlighting that these faces were perceived generally as being neutral. Also, raters were asked to rate the smile, but unable to report a frown.
downward or an upward mouth curvature. A total of 51 new Caucasian (11 non-Caucasian excluded) participants (39 females, 12 males, $M_{\text{age}} = 25.6, SD = 9.44$) were recruited in an online experiment titled ßRate Smiles in Models. ßParticipants again previewed the stimuli (same set of stimuli as above) and were given the following instructions: ßPlease keep in mind that all the faces presented appear to have a neutral expression. None of these faces are meant to have obvious smiles. We are examining very subtle differences in mouth curvature. ßEvaluators were then presented with a face and asked: ßfocusing on the mouth, please rate the degree of mouth curvature in this face compared to the other faces presented. ßEvaluators were able to use an unmarked sliding scale with end points labeled downward mouth curvature and upward mouth curvature; responses on the scale were recorded as an integer from 1 to 100. Perceptual ratings of smiling significantly correlated with objective measurements of mouth curvature ($r(100) = 0.77, p < 0.001$). Results provide evidence that subtle differences in smiling and frowning are perceivable. Indeed perceptual reports differentiate faces with a slight upturn or downturn in mouth curvature and show congruence with image measurements of mouth curvature.

4.3.2 Participants

4.3.2.1 Intelligence evaluators

After exclusions, a total of 173 participants (99 female, 74 male, $M_{\text{age}} = 28.26, SD = 11$) completed an online study entitled ßSmart or Not. ßParticipants assigning very similar ratings to all images (standard deviation of ratings < 0.50) were excluded (13 participants) due to their unusual and limited use of the scale which may indicate inadequate adherence to task instructions or the preconceived notion that social judgments like intelligence cannot be made from facial cues (or is socially undesirable). Further, participants who reported their ethnicity as different from ßwhite Caucasian ß(35) were also excluded, as stimuli presented were Caucasian and judgments of other ethnicities may be more susceptible to stereotypes (Zebrowitz et al., 2007). Participants were asked to

---

4 Non-Caucasian raters were excluded to avoid the ambiguity of any cross-race effects on perceptions of attractiveness and intelligence. Including the non-Caucasian raters for every analysis reported did not change the pattern of results. In most cases, the effects were stronger with these raters included.
rate 100 faces for perceived intelligence. Participants voluntarily gave informed consent and had to be over the age of 18.

4.3.2.2 Attractiveness evaluators

After exclusions, an independent group of 140 participants (69 female, 71 male, $M_{age} = 38.11$, $SD = 10.41$) were recruited for an online study entitled “Adult Attractiveness” using Amazon Mechanical Turk and participated for payment to rate faces for perceived attractiveness. The same exclusion criteria used for intelligence evaluators was used for attractiveness evaluators. Two participants were excluded based on a standard deviation of ratings < .50 and 36 participants who reported their ethnicity as different from “white Caucasian” were also excluded.

4.3.3 Procedure

Participants first completed a questionnaire inquiring about their age, gender, ethnicity, and country of origin. Next, participants previewed all stimuli with each image displayed for one second. The stimuli were then re-presented so that participants could rate the perceived intelligence of each face on a 7-point Likert-type scale with endpoints not at all intelligent and very intelligent. A separate group of evaluators rated the same stimuli on perceived attractiveness of each face on a similar 7-point scale with endpoints not at all attractive and very attractive. In both experiments, faces were presented in a random order.

The minimum viewing time for each image was one second, but no maximum response time was enforced.

4.3.3.1 Statistical analysis

An average score of attractiveness and perceived intelligence was calculated for each of the 100 faces based on the average ratings across all participants. We expected a strong significant correlation between perceived attractiveness and perceived intelligence based on the attractiveness halo. Since our interest was in the impact of facial features on perceived intelligence beyond the attractiveness halo, we used mediation analyses to explore the direct effects of the measurements of eye size, eyelid-openness, and mouth curvature of each face on perceived intelligence and the indirect effects of these variables through perceived attractiveness while controlling for gender. Given that Cunningham’s (1986) measure of eye size is confounded with our measure of eyelid-openness, we expect
these two measurements also to be strongly correlated. To see whether these measures have different relationships with perceptions of attractiveness and intelligence we examine models with the measurements of eyelid-openness and eye size included separately and with both measures in the mediation analysis, as we believe a more valid measure of eye size is one that controls for variation in eyelid-openness.

An exploratory analysis of data was conducted to determine if the variables used were normally distributed. Results for the Kolmogorov-Smirnov test for normality indicated that measurements of mouth curvature and eyelid-openness, as well as averaged ratings of attractiveness and perceived intelligence did not deviate significantly from normal distribution. There were no problems with multicollinearity between variables (Variance Inflation Factor (VIF) for all variables < 1.5).

4.3.4 Results

4.3.4.1 Facial averaging

To help the reader visualize both the nature and subtlety of cues to perceived intelligence, differences are illustrated by facial averages of faces scoring low and high on perceived intelligence (Benson & Perrett, 1992; Benson & Perrett, 1993). Facial averages were constructed to examine the cues of perceived intelligence/attractiveness in male and female faces separately. Facial averages were synthesized from the top 20 male and female faces (10 male and 10 female faces) and bottom 20 male and female faces with the highest and lowest scores on perceived intelligence (Figure 4.3). This process (a) computes the average x and y values for 188 facial landmarks within the set of face images, (b) warps each shape of each facial image into these average coordinates, and then blends the warped component images (Re & Perrett, 2012; Tiddeman et al., 2001). These average images were then made symmetrical to emphasize the role of the cues being investigated (see Perrett et al., 1999).
Figure 4.3. Averaged adult images. High (left) and low (right) perceived intelligence male and female facial averages made symmetrical from Study 1. Objective measures show (slightly) more upturned mouth curvature and eyelid-openness in the high-perceived intelligence facial averages images.
4.3.4.2 Correlation matrix

As expected, there was a strong positive correlation between perceived attractiveness and perceived intelligence ($r(100) = 0.72, p < 0.001$). Although we would expect a high correlation due to the attractiveness halo itself, this can also lead to concerns about the ability to distinguish these constructs empirically. We addressed this in two ways: (1) to correct the correlation for attenuation (see Chapter 2) due to measurement error, and (2) to establish an indicator of discriminant validity. To address the issue of measurement error Cronbach’s Alpha reliabilities were calculated for perceived intelligence ($reliability = 0.98$) and perceived attractiveness ($reliability = 0.98$). After correcting for attenuation, the correlation between attractiveness and intelligence ratings rises from $r = 0.72$ to $r = 0.73$, but this increase is not sufficient to undermine their distinctiveness. To establish discriminant validity the differences in correlations between the two physical measures (mouth curvature and eyelid-openness) and the two perceptual ratings (perceived attractiveness and intelligence) were investigated statistically using the Steiger (1980) test for dependent correlations. The correlation between attractiveness and eyelid-openness was significantly different from the correlation between intelligence and eyelid-openness ($rs = 0.18, 0.33; z = -2.06, p = .039$). Similarly, the correlation between attractiveness and mouth curvature was significantly different from the correlation between intelligence and mouth curvature ($rs = 0.04, 0.24; z = -2.68, p = .007$) supporting discriminant validity based on objective measures from the target faces. Taken together, these two analyses support the distinction between perceived attractiveness and intelligence measures.
As shown in Table 4.1, measurements of eyelid-openness and eye size were strongly correlated. There was no significant correlation between measurements of mouth curvature and measurements of eyelid-openness or measurements of eye size. There were also moderate correlations between gender\(^5\) of face and perceptual ratings, indicating that female faces were perceived as more attractive and intelligent. However, gender was not significantly correlated with measurements of eyelid-openness, eye size, or mouth curvature. Older individuals tended to have upturned mouth curvature and reduced eye size, but there were no significant correlation between age and eyelid-openness, attractiveness, or perceived intelligence.

Table 4.1

<table>
<thead>
<tr>
<th>Adult faces: zero-order correlation matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Perceived Attractiveness</td>
</tr>
<tr>
<td>Perceived Intelligence</td>
</tr>
<tr>
<td>Eyelid-Openness</td>
</tr>
<tr>
<td>Eye Size</td>
</tr>
<tr>
<td>Mouth Curvature</td>
</tr>
</tbody>
</table>

\(*** p < .001. ** p < .01. * p < .05.\)

Note: Sex coded as female=1, male = 0. \(N = 100.\) All tests are two-tailed.

\(^5\) The gender of the face was used as a control throughout the studies because of its influence on perceived intelligence (unlike age). The patterns of findings do not change when not including gender of face as a control.
4.3.4.3 Mediation analysis

To test whether perceived attractiveness mediates the relationship between measurements of eyelid-openness, eye size, mouth curvature, or gender on perceived intelligence\(^6\), the SPSS plugin PROCESS was used (Hayes, 2013). Given the strong correlation between eyelid-openness and eye size, we initially included these in two separate analyses to avoid potential suppression effects. We also included a third mediation analysis incorporating both variables.

The first mediation analysis conducted (see Table 4.2) examined perceived intelligence as the outcome variable, perceived attractiveness as the mediator, and measurements of eyelid-openness and mouth curvature as independent variables, with gender of face as a control variable. More eyelid-openness and female gender significantly predicted perceived attractiveness, but upturned mouth curvature was not a significant predictor (overall model: \(R^2 = 0.17, F(3,96) = 6.47, p < 0.001\)). Perceived attractiveness was a strong predictor of perceived intelligence, yet the direct effects of more eyelid-openness, upturned mouth curvature, and female gender significantly predicted perceived intelligence even with perceived attractiveness in the model (overall model: \(R^2 = 0.63, F(4,95) = 40.01, p < 0.001\)). Bias-corrected confidence intervals for indirect effects were calculated through 5000 bootstrap samples. The indirect effect (i.e. the mediation effect through perceived attractiveness) of measurements of eyelid-openness on perceived intelligence was not significant within the 95% CI (\(B = 2.33, SE = 1.36, 95\% CI [-0.22, 5.14], \beta = 0.13\)), but marginally significant within the 90% CI (\(B = 2.33, SE = 1.36, 90\% CI [0.22, 4.67], \beta = 0.13\)). The indirect effect of measurements of mouth curvature on perceived intelligence through perceived attractiveness was not significant (\(B = 0.60, SE = 0.92, 95\% CI [-1.31, 2.36], \beta = 0.04\)).

---

\(^6\) The variable of age was explored additionally because of the potential effect age can have on eye size and perceptions of naivety. Age was not a significant predictor of perceived intelligence in adults or children (Study 2), and is not included in models for simplicity. Including age in the models does not contribute to the model or change the pattern of results. It is possible that the effects of age on perceptions of intelligence were not significant because the samples in which this potential cue was examined were too narrow in age range.
Table 4.2
Direct and mediated effects of eyelid-openness, mouth curvature, and gender on perceived attractiveness and perceived intelligence. (see Figure 4.4)

<table>
<thead>
<tr>
<th>Criterion / Predictors</th>
<th>B</th>
<th>SE</th>
<th>CI</th>
<th>β</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Attractiveness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyelid-Openness</td>
<td>5.32</td>
<td>2.58</td>
<td>0.21 10.44</td>
<td>0.19</td>
<td>0.042</td>
</tr>
<tr>
<td>Mouth Curvature</td>
<td>1.37</td>
<td>2.21</td>
<td>-3.02 5.76</td>
<td>0.06</td>
<td>0.536</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.56</td>
<td>0.14</td>
<td>-0.84 -0.27</td>
<td>-0.73</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Perceived Intelligence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Attractiveness</td>
<td>0.44</td>
<td>0.05</td>
<td>0.34 0.54</td>
<td>0.61</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Eyelid-Openness</td>
<td>4.54</td>
<td>1.28</td>
<td>2.01 7.08</td>
<td>0.23</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Mouth Curvature</td>
<td>3.73</td>
<td>1.07</td>
<td>1.60 5.86</td>
<td>0.22</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.19</td>
<td>0.74</td>
<td>-0.33 -0.04</td>
<td>-0.34</td>
<td>0.013</td>
</tr>
</tbody>
</table>

Note: The overall model, as well as the indirect and direct effects of mediating variables, are described in-text.

Figure 4.4. Linear regression model for adult faces. This flow chart shows the independent direct effects ($\beta$ values, *p < 0.05) of upturned mouth curvature, eyelid-openness, and attractiveness on perceived intelligence. Gender of face is included in this model, but not displayed (see results for details).
The second mediation analysis (see Table 4.3) examined perceived intelligence as the outcome variable, perceived attractiveness as the mediator, and measurements of eye size (following Cunningham, 1986) and mouth curvature as independent variables, with gender of face as a control variable. Female gender significantly predicted perceived attractiveness, but eye size and upturned mouth curvature were not significant predictors (overall model: $R^2 = 0.16, F(3,96) = 6.12, p < 0.001$). Perceived attractiveness was a strong predictor of perceived intelligence, yet the direct effects of upturned mouth curvature and female gender significantly predicted perceived intelligence even with perceived attractiveness in the model. The direct effects of larger eye size did not significantly predict perceived intelligence (overall model: $R^2 = 0.58, F(4,95) = 32.81, p < 0.001$). The indirect effect of measurements of eye size on perceived intelligence (i.e., the mediation effect through perceived attractiveness) was found to be not significant ($B = 31.21, SE = 18.68, 95\%\ CI [-0.96, 72.44], \beta = 0.12$), nor was the indirect effect of measurements of mouth curvature on perceived intelligence through perceived attractiveness ($B = 0.83, SE = 0.98, 95\%\ CI [-1.16, 2.73], \beta = 0.05$).

Table 4.3
Direct and mediated effects of eye size, mouth curvature, and gender on perceived attractiveness and perceived intelligence.

<table>
<thead>
<tr>
<th>Criterion / Predictors</th>
<th>$B$</th>
<th>$SE$</th>
<th>$CI$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Attractiveness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eye Size</td>
<td>66.92</td>
<td>36.51</td>
<td>-5.55 -139.41</td>
<td>0.17</td>
<td>0.070</td>
</tr>
<tr>
<td>Mouth Curvature</td>
<td>1.77</td>
<td>2.23</td>
<td>-2.66 - 6.20</td>
<td>0.07</td>
<td>0.430</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.53</td>
<td>0.14</td>
<td>-0.82 - 0.25</td>
<td>-0.69</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Perceived Intelligence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Attractiveness</td>
<td>0.47</td>
<td>0.05</td>
<td>0.36 - 0.57</td>
<td>0.65</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Eye Size</td>
<td>13.62</td>
<td>19.02</td>
<td>-24.13 - 51.37</td>
<td>0.05</td>
<td>0.476</td>
</tr>
<tr>
<td>Mouth Curvature</td>
<td>3.76</td>
<td>1.15</td>
<td>1.49 - 6.04</td>
<td>0.22</td>
<td>0.001</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.16</td>
<td>0.08</td>
<td>-0.32 - 0.01</td>
<td>-0.29</td>
<td>0.042</td>
</tr>
</tbody>
</table>

Note: The overall model, as well as the indirect and direct effects of mediating variables, are described in-text.
The third mediation analysis (see Table 4.4) examined perceived intelligence as the outcome variable, perceived attractiveness as the mediator, and measurements of mouth curvature and both eyelid-openness and eye size as independent variables, with gender of face as a control variable. Female gender significantly predicted perceived attractiveness, but upturned mouth curvature, larger eye size and more eyelid-openness were not significant predictors of perceived attractiveness (overall model: $R^2 = 0.17$, $F(4,95) = 4.96$, $p = 0.001$). Perceived attractiveness was a strong predictor of perceived intelligence, yet the direct effect of more eyelid-openness significantly predicted higher ratings of perceived intelligence, while there was a trend for the direct effect of larger eye size to be negatively associated with perceived intelligence with perceived attractiveness in the model. Upturned mouth curvature, and female gender also predicted perceived intelligence with perceived attractiveness in the model (overall model: $R^2 = 0.64$, $F(5,94) = 33.22$, $p < 0.001$). The indirect effect of measurements of eyelid-openness ($B = 1.71$, $SE = 1.67$, 95% CI [-1.53, 5.02], $\beta = 0.08$), eye size ($B = 14.93$, $SE = 20.93$, 95% CI [-24.86, 59.43], $\beta = 0.05$) and mouth curvature ($B = 0.70$, $SE = 0.95$, 95% CI [-1.23, 2.56], $\beta = 0.05$) on perceived intelligence (i.e., the mediation effect through perceived attractiveness) were not significant.

### Table 4.4

<table>
<thead>
<tr>
<th>Criterion / Predictors</th>
<th>B</th>
<th>SE</th>
<th>CI</th>
<th>(\beta)</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Attractiveness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eye Size</td>
<td>33.70</td>
<td>45.95</td>
<td>-57.51 (\tilde{\beta}) 124.91</td>
<td>0.09</td>
<td>0.465</td>
</tr>
<tr>
<td>Eyelid-Openness</td>
<td>3.87</td>
<td>3.26</td>
<td>-2.60 (\tilde{\beta}) 10.34</td>
<td>0.14</td>
<td>0.238</td>
</tr>
<tr>
<td>Mouth Curvature</td>
<td>1.57</td>
<td>2.23</td>
<td>-2.86 (\tilde{\beta}) 6.01</td>
<td>0.07</td>
<td>0.482</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.55</td>
<td>0.14</td>
<td>-0.83 (\tilde{\beta}) -0.26</td>
<td>-0.71</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Perceived Intelligence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Attractiveness</td>
<td>0.44</td>
<td>0.05</td>
<td>0.35 (\tilde{\beta}) 0.54</td>
<td>0.62</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Eye Size</td>
<td>-37.47</td>
<td>22.04</td>
<td>-81.24 (\tilde{\beta}) -6.29</td>
<td>-0.13</td>
<td>0.092</td>
</tr>
<tr>
<td>Eyelid-Openness</td>
<td>6.13</td>
<td>1.57</td>
<td>3.01 (\tilde{\beta}) 9.25</td>
<td>0.31</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Mouth Curvature</td>
<td>3.50</td>
<td>1.07</td>
<td>1.37 (\tilde{\beta}) 5.62</td>
<td>0.20</td>
<td>0.002</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.20</td>
<td>0.07</td>
<td>-0.34 (\tilde{\beta}) -0.05</td>
<td>-0.36</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Note: The overall model, as well as the indirect and direct effects of mediating variables, are described in-text.
4.3.5 Discussion

As expected, there were high correlations between attractiveness and perceived intelligence, reflecting the strength of the attractiveness halo. However, analyses taking into account the attenuation due to measurement error, and correlations demonstrating the discriminant validity of the two constructs support our argument that they can be considered as distinct constructs. This allows us to consider the impact of our other measures of interest on the attractiveness halo effect. Nonetheless, the high correlation does create potential for interpretative difficulties in regression, and under those circumstances we explicitly examine a mediation model that explores their relationship in a way that reflects the traditional understanding of the attractiveness halo. That is, we assumed that perceptions of attractiveness predict perceptions of intelligence, and we examined the direct and indirect effects of our focal and control measures on intelligence, potentially mediated by attractiveness.

The results of Study 1 are consistent with (based on a statistically non-significant trend) Cunningham’s (1986) findings of large eyes being positively correlated with attractiveness. We argue that Cunningham’s measure of eye size is a compound of two variables: the size of the eyes relative to the size of the face and the openness of eyes. When examined conjointly in regression, neither measure of eye shape significantly predicted attractiveness, yet together they did explain a significant portion of variance in ratings of intelligence.

As highlighted by previous research, large eyes are a significant predictor of perceived babyfacedness, which has been shown to cause a baby face overgeneralization effect in which baby-faced individuals are perceived as more naïve (Zebrowitz & Montepare, 1992). When controlling for eyelid-openness and perceived attractiveness in a

7 It may seem surprising that in the regression model neither eye size, nor eyelid-openness was a significant predictor of attractiveness in adults. The lack of significance of each variable as a predictor does not mean that these variables have no influence. The conjoint contribution of the two eye variables is demonstrable in a hierarchical regression model (1st step independent variables: mouth curvature and sex of face, 2nd step independent variables: Cunningham’s eye size and eyelid openness). In the second stage the model showed a significant increase in variance explained by eye size and eyelid openness conjointly ($R^2$ change = 0.08 $p = 0.012$). However, their high correlation means that their unique contributions are undermined when both are included in regressions.
linear regression model predicting intelligence, we found a non-significant trend for large eye size to negatively influence perceptions of intelligence.

In contrast, our new measure of eyelid-openness showed that more eyelid-openness increased perceived attractiveness and intelligence. Further, unlike the possible negative effect of large eye size on perceived intelligence, more eyelid-openness led to higher ratings of intelligence above and beyond the attractiveness halo. Likewise, while faces which had a subtly more positive mouth curvature were not perceived as any more attractive, an upturn in mouth curvature was related to a significant increase in perceived intelligence. Hence, the positive effect evident in the mouth configuration appears to have an impact on the attribution of intelligence independent of an attractiveness halo.

These findings suggest that malleable or dynamic facial features have an impact on social attributions: people whose eyelids were more open were perceived as more intelligent than people who had more drooping eyelids. By the same token, people who had more of an upturned mouth curvature (i.e., showed a subtle smile) were not rated as more attractive, but were rated as more intelligent regardless of their attractiveness.

It should be noted that the faces used in this experiment have been used in other studies of social judgments (e.g., Batres & Perrett, 2014; Quist, DeBruine, Little, & Jones, 2012; Re & Perrett, 2012). In all these experiments, the faces were considered to show “neutral” expressions. The differences in the malleable facial cues of eyelid-openness and upturn in mouth corners need not be based on an overt emotional expression; they could reflect individual differences in anatomy or subtleties in current mood and recent sleep patterns. Subtly upturned mouth curvature and increased eyelid-openness are more visible in the averages of faces judged high on intelligence compared to average of faces judged low on intelligence (see figure 4.3).

Taken together, Study 1 highlights the influence, as well as the subtlety, of the malleable cues of eyelid-openness and mouth curvature on perceptions of intelligence above and beyond the effects of attractiveness. The findings have potential implications in environments in which perceptions of adult intelligence are crucial such as in business, law courts, and higher education. In each of these environments impressions of intelligence may influence expectations about competence. Given the special influence of perceptions of intelligence in education and the particular vulnerability of younger students to the
expectancy effect (Alvidrez & Weinstein, 1999; Rosenthal & Jacobson, 1968), Study 2 set out to replicate findings using children as stimuli.
4.4 Study 2: Children’s Faces

In order to explore the possible impact upon teacher expectancy effects, we investigated the effect of eyelid-openness and mouth curvature on perceptions of intelligence (and attractiveness) in children’s faces. As in Study 1, all faces showed neutral expressions. It was hypothesized that the influence of expressive features on attributions perceived in adult faces would affect children’s faces equivalently. We thus anticipated that eyelid-openness and upturned mouth curvature would benefit attributions of intelligence beyond attractiveness in children’s faces. Successfully replicating the findings in Study 1 across facial images of school-aged children could have significant implications for educational contexts.

4.4.1 Method

4.4.1.1 Face stimuli

Stimuli consisted of children aged 5 to 16 with 49 boys ($M_{age} = 9.38$, $SD = 2.01$) and 41 girls ($M_{age} = 10.19$, $SD = 2.75$). Stimuli pictures were obtained from the Dartmouth Database of Children’s Faces (Dalrymple, Gomez, & Duchaine, 2013). All participants from the database that gave permission to be shared and presented were used as stimuli. All photographs were taken under standardized lighting conditions and camera set-up; individuals wore a black cap, did not wear make-up or jewelry, and posed with a neutral facial expression in front of a black background. As the black cap prevented accurate identification of the hairline, it was not possible to calculate Cunningham’s (1986) measurements of eye size which takes into account eye-height in relation to the length of the face. Delineation, alignment, and cropping of the images followed the procedures outlined in Study 1.

4.4.1.2 Validation of perceptual ratings

A total of 16 new participants (15 female, 1 male, $M_{age} = 23.1$, $SD = 10.9$) were recruited in an online experiment. Seven non-Caucasian participants were excluded but no exclusions were necessary based on standard deviation (all within appropriate range). Participants followed the validation procedures outlined in Study 1. Mouth curvature and
eyelid-openness ratings were assessed in two separate blocks. Block and stimulus order were randomized.

Ratings of mouth curvature and eyelid-openness for each face were averaged across all participants and correlated with the measurements of mouth curvature and measurements of eyelid-openness described in Study 1. The measurements of eyelid-openness and perceptual ratings of eyelid-openness were highly correlated \((r(90) = 0.83, p < 0.001)\). The perceptual ratings of mouth curvature also highly correlated with the objective measurements of mouth curvature \((r(90) = 0.84, p < 0.001)\).

4.4.2 Participants

4.4.2.1 Intelligence evaluators

After exclusions, a total of 76 new participants (45 female, 31 male, \(M_{age} = 34.4, SD = 13.90\)) were recruited for an online study entitled “How Smart Are These Kids.” Participants were asked to rate children’s faces for intelligence. Criteria identical to that in Study 1 were used to exclude participants from the final analysis: three excluded for standard deviation < 0.50 and 14 excluded for not reporting ethnicity as “white Caucasian.”

4.4.2.2 Attractiveness evaluators

After exclusions, an independent group of 60 participants (27 female, 33 male, \(M_{age} = 39.03, SD = 11.26\)) were recruited for an online study entitled “Children Attractiveness” using Amazon Mechanical Turk and participated for payment to rate faces for perceived attractiveness. The same exclusion criteria used for intelligence evaluators was used for attractiveness evaluators. No participants were excluded based on a standard deviation of ratings < .50 and 14 participants who reported their ethnicity as different from “white Caucasian” were excluded.

4.4.3 Procedure

As in Study 1, participants completed a questionnaire inquiring about their age, gender, ethnicity, and country of origin. Further, participants followed the procedures of Study 1 previewing stimuli at one second per face prior to the actual rating. The faces were then re-presented with participants using a 7-point Likert-type scale to rate faces for
intelligence. An independent group of evaluators rated the stimuli for attractiveness using the same endpoints as those in Study 1. In each experiment, faces were presented in a random order with a minimum response time of one second and no maximum response time.

**4.4.3.1 Statistical analysis**

An average score of attractiveness and perceived intelligence was calculated for each of the 90 faces based on the average ratings across all participants. Using similar methods following those used in Study 1, a mediation analysis was conducted to explore the direct and indirect effects of measurements of eyelid-openness and mouth curvature on perceptions of intelligence through, and independent of, attractiveness.

**4.4.4 Results**

**4.4.4.1 Facial averaging**

Again, face averaging processes outlined in Study 1 were used to synthesize 4 facial averages from the top 10 male and 10 female faces and the bottom 10 male and 10 female faces on rated intelligence (Figure 4.5). These composite images were then made symmetrical to make the investigated cues more salient.
Figure 4.5. Averaged children images. High (left) and low (right) perceived intelligence female (bottom row) and male (top row) facial averages made symmetrical from Study 2. Note the slightly more upturned mouth curvature and greater eyelid-openness in the high-perceived intelligence facial average images.
4.4.4.2 Correlation matrix

As expected, and shown in Table 4.5, there was a strong positive correlation between perceived attractiveness and perceived intelligence. There was also a significant correlation between gender and attractiveness such that female faces were perceived as more attractive and intelligent. Gender was not significantly correlated with measurements of eyelid-openness or mouth curvature. Similar to findings from Study 1, there was no significant correlation between age and eyelid-openness, attractiveness, or perceived intelligence. Unlike the findings from Study 1, there was no correlation between age and mouth curvature. There was no significant correlation between measurements of mouth curvature and measurements of eyelid-openness.

Table 4.5
Children faces: zero-order correlation matrix

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Sex</th>
<th>Perceived Attractiveness</th>
<th>Perceived Intelligence</th>
<th>Eyelid-Openness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>-.168</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Attractiveness</td>
<td>-.022</td>
<td>-.276 **</td>
<td></td>
<td>-.276 **</td>
<td></td>
</tr>
<tr>
<td>Perceived Intelligence</td>
<td>-.006</td>
<td>-.268 *</td>
<td>.856 ***</td>
<td>-.268 *</td>
<td></td>
</tr>
<tr>
<td>Eyelid-Openness</td>
<td>-.155</td>
<td>-.057</td>
<td>.248 *</td>
<td>.391 ***</td>
<td>-.311 **</td>
</tr>
<tr>
<td>Mouth Curvature</td>
<td>-.038</td>
<td>-.118</td>
<td>.192</td>
<td>.311 **</td>
<td>.021</td>
</tr>
</tbody>
</table>

Note: Sex coded as female=1, male = 0. N = 90. All tests are two-tailed.

4.4.4.3 Mediation analysis

The mediation analysis (see Table 4.6) examined perceived intelligence as the outcome variable, perceived attractiveness as the mediator, and measurements of eyelid-openness and mouth curvature as independent variables, with gender of face as a control variable. More eyelid-openness and female gender significantly predicted perceived attractiveness, but upturned mouth curvature did not significantly predict perceived attractiveness (overall model: $R^2 = 0.15, F(3,86) = 5.25, p = 0.002$). Perceived attractiveness was a strong predictor of perceived intelligence, yet the direct effects of more eyelid-openness and upturned mouth curvature significantly predicted perceived intelligence with perceived attractiveness in the model. Direct effect of female gender was not significant in predicting perceived intelligence (overall model: $R^2 = 0.79, F(4,85) = $
Bias-corrected confidence intervals for indirect effects were calculated through 5000 bootstrap samples. The indirect effects of measurements of eyelid-openness ($B = 3.96$, $SE = 1.79$, 95% CI [0.50, 7.71], $\beta = 0.19$) on perceived intelligence through perceived attractiveness was significant, but mouth curvature was not ($B = 1.79$, $SE = 1.13$, 95% CI [-0.36, 4.11], $\beta = 0.14$).

Table 4.6
Direct and mediated effects of eyelid-openness, mouth curvature, and gender on perceived attractiveness and perceived intelligence of children’s faces. (see Figure 4.6)

<table>
<thead>
<tr>
<th>Criterion / Predictors</th>
<th>$B$</th>
<th>$SE$</th>
<th>CI</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Attractiveness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyelid-Openness</td>
<td>5.87</td>
<td>2.53</td>
<td>0.85 -10.89</td>
<td>0.23</td>
<td>0.023</td>
</tr>
<tr>
<td>Mouth Curvature</td>
<td>2.65</td>
<td>1.67</td>
<td>-0.67 - 5.98</td>
<td>0.16</td>
<td>0.116</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.31</td>
<td>0.13</td>
<td>-0.56 - 0.06</td>
<td>-0.49</td>
<td>0.017</td>
</tr>
<tr>
<td>Perceived Intelligence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Attractiveness</td>
<td>0.67</td>
<td>0.05</td>
<td>0.58 - 0.77</td>
<td>0.77</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Eyelid-Openness</td>
<td>4.35</td>
<td>1.14</td>
<td>2.09 - 6.60</td>
<td>0.20</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Mouth Curvature</td>
<td>2.29</td>
<td>0.74</td>
<td>0.82 - 3.76</td>
<td>0.16</td>
<td>0.003</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.03</td>
<td>0.06</td>
<td>-0.14 - 0.08</td>
<td>-0.05</td>
<td>0.612</td>
</tr>
</tbody>
</table>

Note: The overall model, as well as the indirect and direct effects of mediating variables, are described in-text.

Figure 4.6. Linear regression model of the perception of children’s faces. This flow chart shows the independent direct effects ($\beta$ values, * $p < 0.05$) of upturned mouth curvature, eyelid-openness, and attractiveness on perceived intelligence. Gender of face is included in this model, but not displayed (see results for details).
4.4.4 Discussion

Results of Study 2 replicated the findings of Study 1 with children’s faces. As in the adult faces, untrained participants were able to identify subtle differences in eyelid-openness and mouth curvature once their attention was drawn to the possible subtlety of the cues. Thus, objectively measured features were found to be congruent with perceptual ratings, validating the measures used. Further, congruent with findings in adult stimuli, upturn in mouth curvature did not significantly influence perceptions of attractiveness, but did predict increased ratings in perceived intelligence, while increased eyelid-openness had beneficial effects on both attractiveness and perceptions of intelligence. Similarly, as in Study 1 and predicted by the attractiveness halo effect, attractiveness ratings were significantly related to intelligence ratings. Hence, attractiveness mediated some of the beneficial effects of these cues on perceptions of intelligence. Yet, eyelid-openness and mouth curvature predicted ratings of intelligence independent of attractiveness. Thus, malleable facial cues to mood and alertness can have both direct and indirect impact on impressions of intellect.

Taken together, Studies 1 and 2 show that in both adults and children more eyelid-openness and upturned mouth curvature result in higher ratings of intelligence regardless of attractiveness. Most studies of facial attractiveness have investigated the impact of facial traits that are fixed (such as symmetry and averageness) or traits that can change only slowly (age, adiposity, skin texture, and color). Our analysis clearly showed that facial cues that are malleable over short timescales could also impact attributions. Yet, because the cues of eyelid-openness and mouth curvature were investigated across different faces, it is possible that other facial features or a pattern of facial cues may be influencing perceptions of intelligence. Study 3 digitally manipulated stimuli to alter eyelid-openness and mouth curvature cues alone and investigated how these cues can independently influence perceptions of intelligence independent of other facial cues.
4.5 Study 3: Malleable Facial Cues Transform

While Studies 1 and 2 found that the cues of eyelid-openness and mouth-curvature influence perceptions of intelligence across different sets of stimuli, this experiment was designed to investigate whether a targeted manipulation of these specific cues was sufficient to influence perceptions of intelligence. In order to isolate eyelid-openness and mouth curvature as independently predicting perceptions of intelligence, it is important to control for other facial features or a constellation of facial cues that may be influencing perceptions. Using stimuli from Study 2, we digitally manipulated the eyelid-openness and mouth curvature independently to examine whether differences in these cues alone would impact perceived intelligence.

4.5.1 Method

4.5.1.1 Facial stimuli

Forty identities (20 female; $M_{age} = 9.90, SD = 2.74$; 20 male, $M_{age} = 9.83, SD = 2.13$) from the Dartmouth Database of Children’s Faces (Dalrymple, Gomez, & Duchaine, 2012) described in Study 2, were selected based on gender and similar age. Using Psychomorph, these 40 identities were transformed (50% in both directions) to create a total of 80 stimuli with differences in eyelid openness and mouth curvature. An average was made of the ten faces (out of the original 90 faces described in Study 2) with the most upturned/downturned mouth curvature and increased/decreased eyelid-openness respectively (irrespective of gender). The mouth shape and eyelid-openness of these averages was used to transform the shape of individual faces. The first 20 randomly selected identities were transformed to create 40 stimuli (half of which were female) with different mouth curvature based on the upturned/downturned mouth curvature averages: 20 with an upturned mouth curvature, 20 with a downturned mouth curvature. A further 20 different identities were also transformed to create 40 stimuli with different eyelid-openness: 20 increased in eyelid-openness and 20 decreased in eyelid-openness (see Figure 4.7).
Figure 4.7. Malleable facial cues transform. An average of all children’s faces was manipulated to have increased eyelid-openness (top left) and upturned mouth curvature (bottom left) and decreased eyelid-openness (top right) and downturned mouth curvature (bottom right). The degree of manipulation for this transform was based on a composite image of faces in the set of stimuli with the most and least eyelid-openness and mouth curvature, respectively.
4.5.2 Participants

After exclusions, a total of 37 new participants (7 males, 30 females; $M_{age}$ 21.65, $SD = 5.12$) were recruited to take part in a study entitled “Influences in the perception of intelligence in faces” in lab study via SONA, an online research participation system for Psychology studies conducted at the University of St. Andrews and open to the community. Participants were compensated for their time at a rate of £5/hour pro rata. The same criteria as those in previous studies were used to exclude participants from the final analysis, with 23 participants excluded for not reporting ethnicity as “white Caucasian” and no exclusions necessary for standard deviation < 0.50.

4.5.3 Procedure

After completing a short questionnaire inquiring about their age, gender, ethnicity, and country of origin, participants rated the perceived intelligence of stimuli using a 7-point Likert-type scale with endpoints not at all intelligent and very intelligent. The 80 stimuli were arranged into two blocks, each with 10 upturned mouth curvature, 10 down turned mouth curvature, 10 increased eyelid-openness, and 10 decreased eyelid-openness. Each block contained 40 unique identities, i.e., no identity was repeated within one block so that each manipulation was done on a different identity and evaluators would not be presented with the same identity manipulated in different ways in the same block. Across the two blocks the only thing in the facial image to change was the upturn/downturn mouth curvature or the increase/decrease in eyelid-openness. Stimuli were standardized identical to that of Study 2: standardized lighting, black background, black caps, no make-up or jewelry, neutral expression and head posture.

4.5.4 Results

A one-way repeated measures ANOVA showed that the difference in intelligence ratings for stimuli with increased eyelid-openness ($M = 4.15$, $SD = 0.60$) and decreased eyelid-openness conditions ($M = 4.08$, $SD = 0.51$) was statistically significant, $F (1,35) = 12.82$, $p = .001$, partial $\eta^2 = .268$. Similarly, a one-way repeated measures ANOVA showed that the difference in intelligence ratings for stimuli with upturned mouth curvature ($M = 3.98$, $SD = 0.63$) and downturned mouth curvature conditions ($M = 3.79$, $SD = 0.60$) was statistically significant, $F (1,35) = 4.48$, $p = .042$, partial $\eta^2 = .113$. There were no
significant interactions for face gender with eyelid-openness \((F(1,35) = .204, p = .654)\) nor with mouth curvature \((F(1,35) = .761, p = .389)\).

### 4.5.5 Discussion

The results of Study 3 emphasize the role of eyelid-openness and mouth curvature influencing perceptions of intelligence independent of any other fixed facial features or combination of facial features. Given that the facial transforms were made from natural variation in neutral resting expressions, it is evident that even the slightest differences in eyelid-openness and mouth curvature can significantly impact social judgments (see Figure 4.7). Study 1 and 2 found that differences in eyelid-openness and mouth curvature across individuals can influence perceptions of intelligence, while Study 3 found that when these cues alone (through digital manipulation) are changed within the same facial image there is an impact on perceived intelligence.

The findings of Study 3 are limited to single images of people and stimuli that have been digitally manipulated to change targeted cues. To confirm whether changes in appearance of real people have an impact on attributions, Study 4 compared different images of the same person and investigated whether a natural change in facial cues can lead to a change in trait attributions. That is, Study 4 attempts to improve the ecological validity and overall generalizability of findings by examining whether natural changes in eyelid-openness and mouth curvature within the same person alter perceived intelligence.
4.6 Study 4: Sleep-Restricted Faces

To investigate the effect of malleable cues to perceived intelligence within the same stimuli, we had participants rate the perceived intelligence and attractiveness of two facial images of the same person taken under separate conditions: after normal sleep and after sleep restriction. We expected that sleep restriction would be associated with negative mood manifested in a more downturned mouth configuration, and increased tiredness manifested in reduced eyelid-openness (Sundelin et al., 2013). We further predicted that changes in these facial features would impact perceptions of intelligence. A set of raters was then asked to choose which face had more eyelid-openness and was smiling more when comparing the two images of the same person taken under the two separate conditions. Hence, while Studies 1 and 2 showed that eyelid-openness and mouth curvature across individuals significantly influence perceptions, Study 4 investigated whether these effects are evident within the same person.

4.6.1 Method

4.6.1.1 Face stimuli participants

The stimuli pictures presented in this experiment came from the Sleep, Cognition, and Health Lab at Karolinska Institutet in Stockholm, Sweden. Specifically, 25 individuals (14 females) were recruited and then photographed in two separate conditions: baseline and sleep restriction. A series of photographs were taken at each occasion with a standardized camera set up. Participants had their hair pulled back, no make-up or jewelry, and were asked to maintain a neutral facial expression. Five to eight representative photographs from each photo shoot were selected after low-quality photos (e.g., participant blinking) had been removed. A final selection of the most representative photo was done by an independent judge unaware of the experimental design or the sleep condition of the person in the photos. The judge was instructed to choose the photo out of the five to eight available pictures that looked the most like the other photos taken at the same time. The individual judge did not know the participants in the photos. Todorov and Porter (2014) highlight significant differences in person impressions within multiple facial photos of the same person due to random variation and discuss how this can influence accuracy of
personality inferences based on faces. Thus, it was important to select the most standardized stimuli.

These photographs were used for ratings in the second part of the study. The stimuli originally consisted of a total of 50 facial images from the two separate conditions (baseline and sleep-restricted), but one female was excluded prior to ratings for not having her hair properly combed back and another two female faces were excluded after ratings because of open mouths or exaggerated smiles (relative to the other stimuli), leaving the total number of stimuli at 44 facial images (11 females, 11 males, $M_{age} = 22.9, SD = 3.33$).

### 4.6.1.2 Sleep-restriction procedures

Participants had a sleep need of seven to nine hours a night with no reported health problems or sleep disturbances. Participants (for stimuli) came into the lab on two separate occasions in a counter-balanced order; after spending at least eight hours a night in bed for two consecutive nights, and after spending no more than four hours a night in bed for two consecutive nights. The lab visits were at least one week apart. Adherence to the protocol was controlled via actigraphs (a small accelerometer worn on the non-dominant arm measuring activity and giving good information about sleep duration), sleep diaries, and text messages sent to the investigator at lights off and waking.

### 4.6.1.3 Validation of objective measures of feature changes

The purpose of this additional validation was to assess whether evaluators were capable of detecting a change in eyelid-openness and mouth curvature between sleep restriction conditions and whether these assessments matched objective measurements of eyelid-openness and mouth curvature. Studies 1 and 2 examined subtle differences in eyelid-openness and mouth curvature between different individuals. Study 4 enabled us to use pairs of facial images of the same individual to examine a more sensitive forced-choice paradigm to explore perceptions of these facial features.

After exclusions, a total of 90 new participants (67 female, 23 male, $M_{age} = 27.3, SD = 11.10$) were recruited in an online study. Thirteen non-Caucasian participants and 22 participants who spent less than one second looking at an image were excluded. Participants completed a forced-choice experiment to compare the eyelid-openness and mouth curvature of the same person, between conditions. Specifically, participants were
shown two images of the same face side-by-side, (baseline and sleep restricted) and were asked to “Please click on the face that is smiling more” or “Please click on the face that has its eyelids more open” in two separate blocks. Each participant completed both blocks (22 trials each block) and the order of blocks and stimulus faces was randomized.

Perceptual evaluations of eyelid-openness and smiling were calculated as the total count of observers choosing one face image over the corresponding face image from the alternative condition. This total count was then divided by the total number of observers (90 participants). This proportion defined perceptual evaluation of each face, which was then compared to the change in measurements of eyelid-openness and mouth curvature between conditions. In a partial correlation controlling for the age and gender of the face, the perceptual evaluation of eyelid-openness was positively correlated with the measured change in eyelid-openness ($r(20) = .62, p = 0.003$) and the perceptual evaluation of smiling was significantly correlated with the measured change in mouth curvature ($r(20) = 0.82, p < 0.001$) across the conditions. Thus, the malleable cues were indeed evident to observers and understood in a way that is consistent with the interpretation given to the objective measures.

4.6.2 Participants

4.6.2.1 Attractiveness and intelligence raters

After exclusions, a total of 61 participants (30 females, 30 males, 1 of unspecified gender, $M_{age} = 25.9, SD = 6.10$) rated 48 facial photographs with respect to attractiveness and intelligence. Participants were recruited at Universities and a workplace in the Stockholm area and received a movie ticket for their participation. Participants who did not speak Swedish, were younger than 18, older than 45, or students of psychology were not recruited. No exclusions were necessary based on the same criteria used to exclude participants in previous studies (i.e., ethnicity and standard deviation of ratings).

4.6.3 Procedure

Participants rated 48 facial images for perceived intelligence and attractiveness. As part of a larger data collection, facial images were also rated for perceived tiredness, leadership ability, trustworthiness, and employability. The first four blocks of ratings were in a randomized order and consisted of intelligence, trustworthiness, leadership ability, and
employability. After that, attractiveness and tiredness were rated, in that order. After a short test trial of four photos to familiarize the participants with the procedure, the photos were presented one-by-one and rated on 7-point Likert-type scales from 1 (not at all intelligent or very unattractive) to 7 (very intelligent or very attractive). The photographs were shown at a self-paced interval, for a maximum of six seconds. The presentations were made in a randomized order, with the exception that the two photographs from the same person (one from each condition) could not be presented back to back. After every 48th photograph, the participants performed a working memory task to prevent memorization of the faces and to avoid a familiarity effect.

**4.6.3.1 Statistical analysis**

The ratings of attractiveness and intelligence for each face were averaged across raters and were then separated by condition. To examine the within-subject differences in eyelid-openness and mouth curvature pre- and post-sleep restriction, the change in ratings for the same face but different condition was computed (averaged baseline ratings subtracted by averaged sleep restricted ratings). This change of ratings between conditions was then compared to the measurements of change (baseline subtracted by sleep restriction) in eyelid-openness and mouth curvature across the 2 conditions. Gender was included as an individual level factor. To remain consistent with previous data analysis strategies, mediation analysis was conducted to examine the direct and indirect effects of the change in measurements of eyelid-openness and mouth curvature on change in perceived intelligence. However, because this is a within subject design with a small sample examining whether a change in malleable facial cues can change perceptions of intelligence within the same person, we were less concerned with whether this effect happened through or independent of the change in attractiveness and instead examined the total effects of these facial cues on the overall change in perceptions of intelligence.
4.6.4 Results

4.6.4.1 Facial averaging

Again, described effects are visualized using the face averaging process explained in Study 1. Facial averages were synthesized from the top 10 male and female (n = 20) and the bottom 10 male and female (n = 20) faces; this time based on change in perceived intelligence within the same individual, rather than average perceived intelligence across individuals (see Figure 4.8).
Figure 4.8. Averaged sleep-restriction images. High (left) and low (right) perceived intelligence male and female facial image averages from Study 4. There seems to be slightly more upturned mouth curvature and eyelid-openness in the high-perceived intelligence (left) facial composite images.
4.6.4.2 Sleep condition

Pre- and post-sleep restriction measurements of eyelid-openness and mouth curvature, along with ratings of attractiveness and perceived intelligence, were compared to investigate whether the condition of sleep restriction influenced any of these variables. Paired-samples t-tests were used to compare measures and ratings across conditions (baseline vs. sleep restricted). For eyelid-openness, there was a significant difference in the scores for baseline ($M = 0.12, SD = 0.02$) and sleep restricted ($M = 0.11, SD = 0.02$) conditions; $t(21) = 3.16, p = 0.005$. The mean difference in eyelid-openness was 0.01 (95% [0.00, 0.01]) demonstrating a medium effect size, $d = 0.67$. For mouth curvature, there was not a significant difference in the scores for baseline ($M = 0.00, SD = 0.03$) and sleep restricted ($M = -0.01, SD = 0.03$) conditions; $t(21) = 1.43, p = 0.169$. For attractiveness, there was a significant difference in the scores for baseline ($M = 3.62, SD = 0.76$) and sleep restricted ($M = 3.48, SD = 0.82$) conditions; $t(21) = 2.70, p = 0.013$. The mean difference in attractiveness was 0.14 (95% [0.25, 0.33]) demonstrating a medium effect size, $d = 0.58$. For perceived intelligence, there was a non-significant trend for a difference in the scores for baseline ($M = 4.46, SD = 0.59$) and sleep restricted ($M = 4.38, SD = 0.61$) conditions; $t(21) = 1.82, p = 0.083$. While the mean ratings did not differ significantly between conditions, it is possible, nonetheless, for changes in eyes and mouth to underlie changes in ratings across the two conditions.
4.6.4.3 Correlation matrix

As expected, there was a strong positive correlation between change in perceived attractiveness and change in perceived intelligence ($r(22) = 0.75, p < 0.001$). There was no significant correlation between gender of face (female coded as 0) and average perceived attractiveness ($r(22) = -0.06, p = 0.781$) or perceived intelligence ($r(22) = -0.13, p = 0.574$). There was also no significant correlation between age of face and perceived attractiveness ($r(22) = 0.07, p = 0.768$) or perceived intelligence ($r(22) = -0.06, p = 0.776$). There was also no significant correlation between change in measurements of mouth curvature and change in measurements of eyelid-openness ($r(22) = 0.10, p = 0.666$).

4.6.4.4 Mediation analysis

The mediation analysis (see Table 4.7) examined change in perceived intelligence as the outcome variable, change in perceived attractiveness as the mediator, and change in measurements of eyelid-openness and measurements of mouth curvature as independent variables, with gender of face as a control variable. Neither gender, change in more eyelid-openness, nor change in upturned mouth curvature were significant predictors of change in perceived attractiveness (overall model: $R^2 = 0.20, F(3,18) = 1.46, p = 0.258$). Change in perceived attractiveness was a strong predictor of change in perceived intelligence, yet the direct effects of change in upturned mouth curvature significantly predicted perceived intelligence with perceived attractiveness in the model. The direct effects of more eyelid-openness and gender were not significant predictors of change in perceived intelligence (overall model: $R^2 = 0.76, F(4,17) = 13.48, p < 0.001$). Bias-corrected confidence intervals for indirect effects were calculated through 5000 bootstrap samples. The indirect effects of change in measurements of eyelid-openness ($B = 2.78, SE = 2.71, 95\% CI [-2.24, 8.75], \beta = 0.19$) and mouth curvature ($B = 1.37, SE = 1.11, 95\% CI [-0.68, 3.78], \beta = 0.19$) on perceived intelligence through perceived attractiveness were both not significant.

However, the total effect of change in more eyelid-openness ($B = 6.34, SE = 3.01, 95\% CI [0.12, 12.67], p = 0.050, \beta = 0.36$) and change in upturned mouth curvature ($B = 4.79, SE = 1.32, 95\% CI [2.02, 7.56], p = 0.002, \beta = 0.60$) on change in perceived intelligence was significant (overall model: $R^2 = 0.51, F(3,18) = 6.33, p = 0.004$).
Table 4.7
Direct and mediated effects of change in eyelid-openness, change in mouth curvature, and gender on changes in perceived attractiveness and perceived intelligence of adult faces with and without sleep restriction. (see Figure 4.9)

<table>
<thead>
<tr>
<th>Criterion / Predictors</th>
<th>B</th>
<th>SE</th>
<th>CI</th>
<th>β</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Attractiveness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyelid-Openness</td>
<td>5.44</td>
<td>4.20</td>
<td>-3.40 - 14.27</td>
<td>0.28</td>
<td>0.212</td>
</tr>
<tr>
<td>Mouth Curvature</td>
<td>2.68</td>
<td>1.84</td>
<td>-1.18 - 6.55</td>
<td>0.31</td>
<td>0.162</td>
</tr>
<tr>
<td>Gender</td>
<td>0.10</td>
<td>0.10</td>
<td>-0.12 - 0.32</td>
<td>0.41</td>
<td>0.356</td>
</tr>
<tr>
<td>Perceived Intelligence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Attractiveness</td>
<td>0.51</td>
<td>0.12</td>
<td>0.25 - 0.77</td>
<td>0.55</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Eyelid-Openness</td>
<td>3.57</td>
<td>2.27</td>
<td>-1.23 - 8.37</td>
<td>0.20</td>
<td>0.135</td>
</tr>
<tr>
<td>Mouth Curvature</td>
<td>3.42</td>
<td>1.01</td>
<td>1.30 - 5.54</td>
<td>0.43</td>
<td>0.003</td>
</tr>
<tr>
<td>Gender</td>
<td>0.01</td>
<td>0.06</td>
<td>-0.11 - 0.13</td>
<td>0.05</td>
<td>0.846</td>
</tr>
</tbody>
</table>

Note: The overall model, as well as the indirect and direct effects of mediating variables, are described in-text.

Figure 4.9. Linear regression model of the perceptual effects of sleep restriction. This flow chart shows the standardized independent direct effects (β values, * p < 0.05) of change in mouth curvature and change in eyelid-openness on change in perceived attractiveness and intelligence. Gender of face is included in this model, but not displayed (see results for details).
4.6.5 Discussion

Study 4 focused on how malleable cues influence perceived attractiveness and intelligence in the same person rather than across different individuals (Studies 1 and 2). The findings suggest that the malleable cues of smiling and eyelid-openness do influence perceptions of intelligence in the same person regardless of the effect of sleep condition on these cues. Interestingly, there was no significant relationship between change in eyelid-openness and smiling on perceived attractiveness, yet both malleable facial features had an influence on the change in perceived intelligence. The importance of smiling perception is apparent from the finding that images with a more upturned mouth curvature were perceived as more intelligent, independent of the impact of the attractiveness halo. Hence, regardless of fixed facial traits (such as averageness, symmetry, gender, and age), malleable cues can alter perceptions of intelligence.

Study 4 also examined the role of sleep restriction on attributions to faces and malleable facial features. It was found that sleep restriction was associated with people having less eyelid-openness, but sleep restriction was not associated with a change in mouth curvature. This gives objective support to the finding that sleep deprived people are perceived as having more heavy eyelids, but contrasts the finding that sleep deprived people are also judged to have more droopy mouth corners (Sundelin et al., 2013). We note that the previous study on this had a more severe form of sleep deprivation, and the degree of sleep deprivation may influence the effects on mouth curvature. Regardless, the change in mouth curvature from one photographic condition to another was significantly associated with perceptions of intelligence. That is, when comparing two images of the same person, the image with the more positive mouth curvature looks more intelligent. Thus, smiling is such a controllable cue that even people who are sleep deprived can be perceived as more intelligent (compared to a rested state) if they exhibit the subtlest upturn in mouth curvature.

Taken together, the findings of Study 4 suggest eyelid-openness and mouth-curvature can subtly change within the same person and that these changes are perceivable and influence attributions of intelligence accordingly. Specifically, when a face had more eyelid-openness and more upturned mouth curvature, that face was perceived as more intelligent compared to an alternative facial image of the same person with less eyelid-
openness and more downturned mouth curvature. Some fixed facial attributes such as averageness, sexual dimorphism, and skin pigment evenness affect facial attractiveness and may affect attribution of traits including intelligence. We have argued that malleable facial cues also have a role in attributions, but the degree of eyelid-openness and mouth curvature can also differ anatomically between individuals; one individual may on average have more droopy eyelids than another individual independent of the sleep habits of the two. All of these fixed traits may have a role in attributions made from facial appearance and could contribute to ratings in Studies 1 and 2. Studies 3 and 4 provide much more direct evidence that facial features changing within an individual can have effects on attributions of intelligence.

4.7 General Discussion

The aim of the presented studies was to introduce and validate new measures to investigate malleable cues to perceptions of intelligence that are independent of the attractiveness halo. The cues of eyelid-openness and mouth curvature were targeted because of their likely relationship to sleep-deprivation and low mood, both of which have been associated with decreased cognitive ability (Ellis et al., 1984; Thomas et al., 2000). Our findings supported the hypothesis that more eyelid-openness and upturned mouth curvature increase ratings of perceived intelligence independent of attractiveness across adult faces (Study 1), children’s faces (Study 2), and even within the same individual’s face (Studies 3 and 4). We have focused on generalizability of the effect across ages and within the same person, and while we expect that the general point we are making would apply to other-race faces, we have limited our sample here for control purposes. The generalizability of the effects across races will need to be established in the future with a broader sample of faces and raters.

4.7.1 Attractiveness

Given the strength and pervasiveness of the attractiveness halo, it was important to determine whether the effects of the investigated cues stem from overgeneralizations of conditions that limit cognitive performance (i.e., alertness and mood) or only influence perceptions of intelligence through their impact on perceived attractiveness. It is clear from our findings that while the attractiveness halo is robust, there are certain malleable facial
cues that can be changed to make it more likely that someone will be perceived as more or less intelligent regardless of their physical attractiveness.

4.7.2 Overgeneralizations

The beneficial impact of eyelid-openness and upturned mouth on perception of intelligence are likely to reflect overgeneralizations. Tiredness (Thomas et al., 2000; Walker, 2009) and low mood (Ellis et al., 1984) detract from current cognitive ability. Evaluators may therefore generalize from cues to tiredness and poor mood to the trait of intelligence; someone who looks awake and happy is thought to be more intelligent than someone who looks tired and unhappy. Overgeneralizations from mouth shape to emotional state has been previously suggested to underlie social attributions (Oosterhof & Todorov, 2008). Our findings suggest overgeneralizations arise from the eyes as well as the mouth. Moreover, we show that both face cues are detected by observers and can be objectively measured in images.

4.7.3 Mouth curvature

The shape of the mouth (expressing a subtle smile or frown) is a malleable facial cue that does not necessarily seem to influence perceived attractiveness across adult faces (Study 1) but does increase attractiveness in children’s faces (Study 2) and significantly influences perceptions of intelligence in both adults (Study 1) and children (Study 2). It was also shown that a positive change in mouth curvature (more smiling) results in an increase in perceived intelligence ratings independent of other facial cues (Study 3) and within the same person (Study 4).

While the mouth curvature is so subtle that the facial expressions seem neutral (see Figures 3, 4, and 5), our findings suggest that observers are able to detect these subtle differences in mouth curvature and that these differences influence perceptions of intelligence accordingly. The reason smiling influences perceptions of intelligence above the attractiveness halo may be an overgeneralization effect related to sad mood and its effect in reducing cognitive performance (Ellis et al., 1984) or subtle frowning may indicate tiredness, which is related to worse cognitive functioning (Thomas et al., 2000; Walker, 2009).
More generally however, the finding that subtle cues may impact trait ratings has important implications. Oosterhof and Todorov (2008) argue that impressions of emotional valence in “neutral” faces can affect trait judgments. While the stimuli faces in the current study did not seem to be smiling overtly or noticeably to the participants, a positive mouth curvature could enhance judgments of attractiveness and perceived intelligence and conversely subtle frowns could have adverse effects on trait evaluation. Further, it could be argued that unintentional smiles are a reflection of mood (i.e., a better mood might be accompanied by a subtle smile). Based on findings that a sad mood is associated with reduced cognitive performance (Ellis et al., 1984), signs of mood could potentially be a reflection of current cognitive capacity.

These results are important to consider in future studies where experimenters may assume facial stimuli to be “neutral.” Differences in mouth curvature, which observers do not report as differences in smiling, seem to affect perceptions of intelligence. Indeed, measurements of mouth curvature may reveal effects on a variety of different social judgments of faces (e.g., perceived dominance, trustworthiness, etc.).

4.7.4 Eyelid-openness

Eyelid-openness is another malleable facial cue that influences perceived intelligence independent of the attractiveness halo. The impact of eyelid-openness was observed both across faces of different people (Study 1 and 2) and within the face of the same person (Study 3 and 4). Observers validated the measurements used to calculate eyelid-openness variation across the faces of different people (adults and children) and within the face of the same person across two different conditions (Study 4).

It could be argued that although eyelid-openness is not fixed, one’s ability to change eyelid-openness is less controllable than one’s ability to control mouth shape. While attempting to increase eyelid-openness (to look more alert or intelligent) is possible, it is difficult to do so without moving the eyebrows and looking surprised or unnatural. Perhaps unsurprisingly, results indicated that sleep restriction reduced eyelid-openness which was in turn detrimental to ratings of perceived intelligence.

We also found that in line with Cunningham’s (1986) earlier findings, measurement of eye size correlated positively with ratings of attractiveness. However, when controlling for attractiveness and eyelid-openness, the neotenous fixed trait of large eyes is perceived.
as less intelligent, congruent with the baby-faced overgeneralization effect. Hence, while large eyes may positively influence attractiveness, large eyes relative to the size of the face can decrease rather than augment perceived intelligence. Further, the difference in direction between the effects of eyelid-openness and eye size on perceptions of attractiveness and intelligence points to the distinction between the perceived attractiveness and intelligence constructs despite their high correlations with each other.

4.7.5 Gender

Gender of the face was included as a factor in the analyses in order to control for any effects it may have on eye size and independent influences it may have on perceptions of intelligence. In adults (Study 1) and children (Study 2) faces, gender of the face significantly influenced perceptions of intelligence independent of attractiveness; females were judged to look more intelligent than males. These findings conflict with previous literature that finds females are perceived as less intelligent than males (Beloff, 1992; Byrd & Stacey, 1993). Specifically, Furnham and Rawles (1995) found that both men and women rate their fathers as having higher intelligence than their mothers and rated their grandfather's intelligence higher than their grandmother's intelligence. Additionally, previous research has also highlighted that parents think their sons are brighter than their daughters (Furnham et al., 2002).

Such perceptions may be outdated as there is increasing evidence that women do better academically than men (Gurian & Stevens, 2004). Gender was primarily included in our analyses as a control to the effects on the facial cues investigated. More research into the characteristics of the raters is needed to establish whether our findings represent a potential societal shift in perceptions of women's intelligence relative to men.

4.8 Conclusions and Implications

We have documented two malleable facial cues that influence first impressions and may create an expectancy effect. We find that there is a strong relationship between the cue of eyelid-openness and perceptions of both attractiveness and intelligence, yet this cue has been largely overlooked in literature on social perception of faces. Further, we find mouth curvature in faces with a neutral expression have a pronounced influence on perceptions of
intelligence. We show the effects of mouth curvature and eyelid-openness on facial perception is widespread, influencing intelligence judgments in adults and children alike.

The judgments of attractiveness and competence have a broad range of implications influencing perceptions of success (Rule & Ambady, 2010), hiring and promotion (Rule & Ambady, 2009), and mate selection (Eagly et al., 1991; Moore et al., 2011). The implications for children are perhaps most important. Children arriving at school tired from inadequate sleep or in low mood due to adverse home life are likely to be judged as less intellectually able. In turn these judgments are likely to lower teacher expectations and hinder educational attainment (Rosenthal & Jacobson, 1968). Yet, we show that the cues on which intelligence judgments hinge are changeable. Hence, improvements in sleep patterns and mood for any individual may benefit social evaluations and expectations. The subtlety of the cues studied here means that the role of malleable facial features in society and their influence in diverse situations, such as in the formation of teacher expectations and employee recruitment, is likely to have been underestimated.
Chapter 5: Own Intelligence and Endorsement of the Intelligence-Attractiveness Halo

This chapter is based on research that is under peer-review in an academic journal:

5. Abstract

While some theories emphasize the influence of the "attractiveness halo" on perceptions of intelligence, there is also evidence which suggests that perceptions of attractiveness themselves can be influenced by perceptions of other desired traits such as intelligence. In an educational context, the effect of impressions of intelligence on teachers' expectations of students gives them particular significance. Research on kin selection and cognitive biases highlights the possibility that intelligent people endorse the intelligence-attractiveness relationship more strongly than less intelligent people. We investigated how a perceiver's own intelligence can influence the association between perceived intelligence and attractiveness of others. We asked 126 participants to rate 48 children's faces for perceived intelligence and attractiveness and then asked them to complete the International Cognitive Ability Resource (ICAR) intelligence test. Ratings by participants who scored higher on the intelligence test showed a stronger relationship between perceptions of intelligence and attractiveness than participants who scored lower on the intelligence test. This effect was significant even after controlling for differences in participants' scale use, i.e., controlling for \( SD \) and \( M \) of ratings of attractiveness and perceived intelligence. These findings, while preliminary, illuminate an individual difference that influences perceptions of intelligence with potentially concerning implications regarding expectancy effects in educational settings.
5.1 Introduction

The widely studied halo effect suggests that certain traits function as a metaphorical halo, casting an overly positive light on other traits. Thorndike (1920) defined the halo effect as a tendency to form a general evaluation of someone as good or bad and to base future judgments of a person based on this general feeling. In a comprehensive analysis of the halo effect, Asch (1946) asserted that impression formation of individuals involves a holistic process of attempting to form an impression of the entire person, based on dynamic interactions of various traits, rather than isolated traits forming the impression of a part of a person. In turn, the halo effect can lead to general impression formation, centered on insufficient or limited information relying on isolated traits.

The halo effect has been studied extensively in the context of education because of the influence that general impressions may have on expectations of students (Rosenthal & Jacobson, 1968) and the consequences of expectancy effects on student performance (de Boer, Bosker, & van der Werf, 2010; see Chapter 3). This study further explores the halo effect by investigating the influence that one’s own intelligence may have on the degree to which an individual endorses the intelligence-attractiveness relationship in facial images of children. It is useful to recognize the various potential origins of the intelligence-attractiveness halo to understand the potential role of own intelligence as an individual difference related to the endorsement of the intelligence-attractiveness association.

5.1.1 Individual differences

The difference in an individual’s inclination to rate a child’s face that is perceived to be intelligent as attractive can be interpreted as either: being more susceptible to the attractiveness halo or having a stronger preference (reflected in higher ratings of attractiveness) for intelligent looking faces. Many studies address the question of attractiveness in the context of theories of assortative mating and mate value (see Chapter 3). Yet, a number of the findings can be interpreted more broadly in terms of preferences for similar others. We therefore briefly consider these findings with that broader interpretation in mind. Carter and Glick (1976) suggested that couples select for similarity and argued against the idea that opposites attract. The tendency to seek out those similar to oneself in regards to intelligence (Jensen, 1978; Price & Vandenberg, 1980; Reynolds et
al., 2000; Vandenberg, 1972) and educational achievement (Mare, 1991) has yielded consistently positive results. Furthermore, significant positive correlations between spouses’ mental abilities have been found to exist even after controlling for socioeconomic status and education (Watkins & Meredith, 1981).

Reciprocally, research has also shown that individual preferences for certain personality traits in others can influence ratings of facial attractiveness (Little, Burt, & Perrett, 2006). That is, if individuals prefer a particular trait, they might consider faces they perceive as having that trait as more attractive than other faces (Little et al., 2006). For example, a person who admires the importance of hard work may find a face that they deem to look hard working as being more attractive. Further, while previous research suggests people may estimate personality from faces with some accuracy (Penton-Voak, Pound, Little, & Perrett, 2006), Little et al., (2006) found that perceptions of personality alone can influence attractiveness and that people who consider a particular personality preferable will have different perceptions of attractiveness based on that liking. Such differences in the perception of preferred traits influencing attractiveness may extend to the perception of intelligence, such that those who are more intelligent value intelligence more as a trait and thus perceive faces that are intelligent as more attractive.

Indeed, it has been argued that, given sufficient time, people are more likely to rely on relevant information about personality or ability to form impressions of others, rather than attractiveness (Campbell, 1979; Eagly et al., 1991; Felson & Bohrnstedt, 1979; Gross & Crofton, 1977; Owens & Ford, 1978). Further, previous research has found that non-physical factors (e.g., information about personality, previous academic achievement) can have a significant influence on perceptions of attractiveness (Kniffin & Wilson, 2004; Zhang et al., 2014). The current study did not examine the direction of the perceived intelligence-attractiveness relationship or the influence of context (i.e., information about the perceived stimuli) on attributions, but focused on individual differences that may be associated with a stronger tendency to rate a child perceived as intelligent also as more attractive. We explore the theories of kin selection and anchoring effects that may explain the potential for individuals who score higher on an intelligence test to find faces of children perceived to be intelligent as more attractive than children perceived to be unintelligent.
5.1.2 Kinship

While assortative mating highlights the tendency for people to choose mates based on similarities, kin selection proposes that individuals will help others in a manner proportionate to genetic similarity (Hamilton, 1964). Nepotism would suggest we favor individuals most similar to us because of the likelihood of kinship (Alexander, 1979). Phenotype matching is one mechanism of kin recognition; an organism can learn their own phenotype and/or those of their familiar kin and later use this learned template in matching the self with the phenotypes of unknown individuals (Holmes & Sherman, 1982; Lacy & Sherman, 1983). Indeed, DeBruine (2002) found that people are more altruistic towards self-resentling individuals, even when this resemblance is very subtle. While Bressan and Martello (2002) found that similar looking individuals are often considered more likely to be genetically related than dissimilar looking people, they also found that belief in genetic relatedness (compared to actual genetic relatedness) was a stronger predictor of perceived similarity. While facial similarity is one mechanism of phenotype matching, belief in genetic relatedness may also stem from similarity on other heritable traits, like intelligence.

It might be considered surprising that similarity would be attractive if it is a cue to kinship, since people are generally averse to sexual relations with kin. A closer examination of the similarity-attraction effect reveals that similarity does not necessarily imply sexual attraction, but rather the liking of another person (Park, Schaller, & Van Vugt, 2008). Thus, people who score higher on intelligence tests may find perceived intelligence more attractive because of a similarity in intelligence (a potential cue to kinship that has shown to influence ratings of likability or attractiveness; Byrne et al., 1968; Byrne, 1961; Byrne, Clore, & Smeaton, 1986; Byrne & Nelson, 1965).

5.1.3 Anchoring effect

The anchoring effect describes the tendency to make decisions that are biased toward the initial judgment (Tversky & Kahneman, 1974). Essentially, the anchoring effect suggests that individuals get stuck on initial attributions when no other information is available (i.e., when rating perceived intelligence from just a face). While some may reason that more intelligent people would be less susceptible to cognitive biases like anchoring effects, a thorough review by Stanovich and West (2008) found that various
cognitive biases (including anchoring effects) are unrelated to cognitive ability. A targeted attempt to investigate individual differences influencing performance on anchoring tasks, namely personality and intelligence, also failed to replicate any benefits of cognitive ability on susceptibility to anchoring effects (Furnham, Boo, & McClelland, 2012).

Conversely, Kahneman and Frederick (2002) argued that while high intelligence respondents have the resources to assist in overcoming easy or typical mistaken intuitions, the authors also argue that when problems become more difficult, the correlation (between intelligence and cognitive bias) is likely to reverse because the more intelligent respondents are more likely to agree on a plausible error than to respond randomly (p.14). Thus, the improved ability to make logical connections and rationalize may actually prove counterproductive to overcoming cognitive biases. Taylor (1923) concurs generally that intelligence is not always a protection against rationalization. Indeed intelligence is what makes rationalization possible (p. 415). Thus, people who score higher on intelligence tests may be no less or even slightly more susceptible to cognitive biases such as the intelligence-attractiveness halo.

5.2 Research Questions

Various theories may account for a strong relationship between perceptions of intelligence and attractiveness. Further to this, the strength in the endorsement of the intelligence-attractiveness relationship may depend on individual differences that stem from a general self-similar bias (e.g., related to assortative mating, mate value, or kin selection). Based on literature suggesting cues to kinship can influence liking of others (Byrne et al., 1968; Byrne, 1961; Byrne, Clore, & Smeaton, 1986; Byrne & Nelson, 1965) and high intelligence does not limit susceptibility to anchoring bias (Stanovich & West, 2008) and may actually exacerbate it (Kahneman & Frederick, 2002), we predicted that individuals who score higher on intelligence tests would be more likely to rate faces they perceive as intelligent also as attractive.

To investigate the degree to which one’s own intelligence influences the extent of a perceived intelligence-attractiveness bias, participants were asked to rate children’s faces for perceived intelligence and then attractiveness. Given the strength of the attractiveness halo’s influence on perceptions of intelligence when no other information is provided, it is
likely participants are at least implicitly using cues related to attractiveness to make judgments of intelligence. Thus, we always obtained ratings of intelligence first and attractiveness second to measure the degree in which individuals find facial cues to intelligence related to attractiveness. After rating stimuli, participants completed a short intelligence test. Monin and Oppenheimer (2005) highlight the important distinction between the correlations of averaged ratings versus averaging individual correlations. Thus, a correlation between perceived attractiveness and perceived intelligence will be calculated for each participant resulting in individual pieces of data that can be compared to an individual’s intelligence score. Each individual’s ratings of intelligence were also compared to the average attractiveness to create two separate halo metrics: correlation between individual intelligence ratings and individual attractiveness ratings across all faces and correlation between individual intelligence ratings and averaged ratings of attractiveness within the sample.

5.3 Method

5.3.1 Facial stimuli

Stimuli consisted of 48\(^1\) photographs of children aged 5 to 16, 24 of which were boys \((M = 10.08, SD = 1.79)\) and 24 of which were girls \((M = 9.87, SD = 3.10)\). Stimuli were obtained from the Dartmouth Database of Children’s Faces (Dalrymple et al., 2013). All photographs were taken under standardized lighting conditions and camera set-up; individuals wore a black cap, did not wear make-up or jewelry, and posed with a neutral facial expression and head posture in front of a black background. Face images were aligned in size and position based on left and right pupils. Images were then resized and cropped \((1608 \times 2584 \text{ pixels})\) so that an equal proportion of head and neck was shown in each image.

---

\(^1\) Forty-eight facial images were randomly selected from the original 90 facial images used in Chapter 4 (Study 2) to create a subset of stimuli for this experiment. Fewer stimuli were used to minimize participant fatigue, as participants were asked to rate both intelligence and attractiveness (rather than just one block of ratings as done in Chapter 4) and do a cognitive ability test.
5.4 Participants

After exclusions\(^2\), a total of 126 participants (47 males, 79 females) aged 18 to 30 (\(M_{age} = 21.12\) years, \(SD = 2.46\) years) were recruited to take part in a study entitled ‘Influences in the perception of intelligence in faces in lab study.’ Participants were recruited through the university’s subject pool consisting of psychology students and community members. Participants were compensated for their time at a rate of £5/hour pro rata. Following similar methods as those used in Chapter 4, participants who reported their ethnicity as different from ‘white Caucasian’ (21) were also excluded, as stimuli presented were Caucasian and judgments of other ethnicities may be more susceptible to stereotypes (Zebrowitz et al., 2007).

5.5 Procedure

5.5.1 Face-rating task

Participants first previewed all stimuli with each image displayed for one second. The stimuli were then re-presented so that participants could rate the perceived intelligence of each face on a 7-point Likert-type scale with endpoints not at all intelligent to very intelligent. The stimuli were then presented a third time to rate the attractiveness of each face on a similar 7-point scale with endpoints not at all attractive to very attractive. Participants had no indication that they would rate the same faces for attractiveness after ratings for intelligence. Since participants were briefed that they would make intelligence judgments, it was felt best to have them rate attractiveness only after completing all intelligence ratings. This minimized any priming effects which may have suggested that attractiveness and intelligence could be linked or that we were measuring the link between the two ratings. In both blocks, faces were presented in a random order. The minimum viewing time for each image was one second, but no maximum response time was enforced.

\(^2\) The recruited sample was not restricted to university students, but the majority of participants were university students between the ages of 18 and 29. Four participants (two 34-year-olds, one 41 year-old, and one 56 year old) were excluded from the analysis as outliers in age. Re-running the analysis with these participants did not change the pattern of findings.
5.5.2 Intelligence measure

After rating the perceived intelligence and attractiveness of each image, participants took a short intelligence test drawn from the International Cognitive Ability Resources (ICAR) test (see Appendix 1), which has been shown to be moderately to strongly correlated with measures of cognitive ability and achievement (Condon & Revelle, 2014). The ICAR assessment is administered online and is an untimed assessment consisting of 16 items divided into four item types: verbal reasoning, letter and number series, matrix reasoning, and three-dimensional rotation. An average score was calculated for every participant based on the number of questions answered correctly, so that the score therefore represents the proportion of the test answered correctly (out of 1).

5.5.3 Statistical analysis

At the stimuli level, averaged ratings of attractiveness and intelligence were calculated for each face (48 total faces) based on the 126 raters. We compared averaged ratings of attractiveness and perceived intelligence, expecting a strong correlation based on the "halo effect." Averaged ratings of attractiveness and perceived intelligence were also compared to the age and gender of the face.

Using similar methods as those used by Monin and Oppenheimer (2005), we computed correlations at the individual level. A correlation between participants' ratings of intelligence and ratings of attractiveness for all 48 faces was calculated for each individual participant; the correlation value is hereafter named the "individual halo." The stronger this correlation, the more an individual's ratings for perceived intelligence resembled their ratings of perceived attractiveness for the facial stimuli presented (see Figure 5.1). While we expected a strong correlation between perceived attractiveness and perceived intelligence based on the halo effect, we investigated whether individuals' intelligence (measured by ICAR) could predict differences in the halo (i.e., the strength of the perceived intelligence and perceived attractiveness correlation across 48 faces).
Figure 5.1. Variation in strength of the individual intelligence-attractiveness halo across individuals. Scatter plots visualizing the association of perceived intelligence and perceived attractiveness for an average of the 5 individuals with lowest attractiveness-intelligence halo (left, $r(48) = .25$) and highest (right, $r(48) = .94$) on the halo metric. Both perceived intelligence and attractiveness were rated on a 7 point Likert scale.

Given the need to get the same participant to rate perceived intelligence and attractiveness of the same stimuli, it was possible that one task would influence ratings on the next task. We believed having individuals rate intelligence first would mitigate this influence, yet we realize this does not eliminate potential bias. Thus, we calculated a separate 'halo' metric between individual ratings of intelligence and the average attractiveness ratings (averaged across all 126 participants in the sample); the correlation value for this metric is hereafter named the 'average halo'\textsuperscript{3}.

A Kolmogorov-Smirnov test showed that measurements (ICAR) of intelligence ($p = .091$) and the perceived intelligence and attractiveness correlation ($p = .200$) across subjects were normally distributed. There were no problems with multicollinearity between

\textsuperscript{3}A third halo metric was also calculated in which individual ratings of intelligence were compared to independent ratings of attractiveness (those collected in Chapter 4, Study 2) and can be referred to as ‘independent halo’. There was no relationship between ‘independent halo’ and one’s own intelligence ($r(126) = .104$, $p = .248$). The relationship between the ‘independent halo’ was strongly correlated ($r(126) = .98$, $p < .001$) with ‘average halo’; given their near indistinguishable correlation, only one metric is discussed for simplicity.
these two variables (Variance Inflation Factor for all variables < 1.5). The “average halo” metric (correlation between individual ratings of intelligence and average ratings of attractiveness) was not normally distributed \( p = .028 \) so non-parametric statistics were used to analyze these data accordingly.

Additionally, the standard deviation (SD) and mean (M) rating of intelligence and attractiveness were calculated for each participant based on their ratings across the 48 faces. These measures of scale use provided a control for how they may influence the intelligence-attractiveness correlation halo. While the M and SD cannot influence the intra-individual correlation between attractiveness and intelligence, because Pearson’s \( r \) correlation already controls for them, the differences in scale use between individuals could have an effect on group level analysis. For this reason, the SD and M ratings of attractiveness and intelligence were used in linear regression to ensure any effects observed would account for differences in scale use.

5.6 Results

As expected, a by-stimulus analysis found that averaged ratings of attractiveness and perceived intelligence were strongly correlated \( r(48) = 0.95, p < .001 \). There was no significant correlations between gender of face (female coded as 0, male coded as 1) and perceived attractiveness \( r(48) = -0.24, p = .105 \) or perceived intelligence \( r(48) = -0.21, p = .146 \). Similarly, there was no significant correlation between age of the individual depicted in the facial stimuli and perceived attractiveness \( r(48) = -0.03, p = .854 \), nor between age and perceived intelligence \( r(48) = 0.01, p = .962 \).
5.6.1 Correlation matrix

At the individual level, the zero-order correlation matrix presented in Table 5.1
shows that there was a significant positive correlation between the intelligence-
attractiveness rating correlation "halo" calculated for each participant and the participants' IQ score ($r(126) = 0.21, p = .018$) (see Figure 5.2). Variables related to scale use ($SD$ and $M$ of perceived attractiveness and intelligence ratings) were not significantly correlated to participants' intelligence-attractiveness halo. Participants' intelligence-attractiveness halo was also not correlated with the participant age or gender. Participants' mean attractiveness and intelligence ratings were strongly correlated ($p < 0.001$). This reflects participants varying in generosity with both judgments (i.e., if a participant generally gave high ratings of attractiveness across all stimuli, they did so similarly when rating intelligence).

Similarly, the participants' standard deviation of ratings for attractiveness and intelligence were strongly correlated ($p < .001$). This indicates the tendency to use the full scale range (i.e., 1-7) or to confine scores to a more restricted range (i.e., 2-4) in both judgments and that participant were using the scale similarly for both ratings of attractiveness and intelligence.

<table>
<thead>
<tr>
<th>Table 5.1</th>
<th>Zero-order correlation matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Halo</td>
</tr>
<tr>
<td>Participant Age</td>
<td>.013</td>
</tr>
<tr>
<td>Participant Sex</td>
<td>.033</td>
</tr>
<tr>
<td>Participant IQ Score</td>
<td>.210*</td>
</tr>
<tr>
<td>$SD$ Attractiveness</td>
<td>.130</td>
</tr>
<tr>
<td>$M$ Attractiveness</td>
<td>.043</td>
</tr>
<tr>
<td>$SD$ Intelligence</td>
<td>.111</td>
</tr>
<tr>
<td>$M$ Intelligence</td>
<td>-.078A</td>
</tr>
</tbody>
</table>

*p < .001.  **p < .01.  *p < .05,  †p < .1. Two-tailed probabilities.

Note: Sex is coded female = 0, male = 1. Correlations are based on 126 participants. "Halo" is the correlation for participants' ratings of intelligence and attractiveness.
Figure 5.2. Participants’ actual intelligence and halo scatter plot. This scatter plot shows the positive correlation between participants’ actual intelligence (1 = every question answered correctly) and the strength of their intelligence-attractiveness halo (or correlation).
There was no significant correlation (Spearman’s rho⁴) between the ‘average halo’ and individual’s own intelligence ($r_s = 0.06, p = 0.504$), $M$ of intelligence ratings ($r_s = -0.08, p = 0.389$), or $SD$ of intelligence ratings ($r_s = 0.11, p = 0.210$).

### 5.6.2 Regression

A multi-step hierarchical linear regression model was conducted to investigate whether participants’ IQ score would predict the strength of their intelligence-attractiveness halo (Table 5.2) with and without control for scale use variables in the model. In step 1 of the model, the participant’s IQ score was entered as a predictor of the intelligence-attractiveness halo. In step 2 of the model, scale use variables ($SD$ and $M$ of attractiveness and intelligence ratings) were entered to account for any differences in scale use. Both models 1 and 2 indicated that participants’ IQ score significantly predicted ($p = .018$) participants’ intelligence-attractiveness halo (see Table 5.2).

<table>
<thead>
<tr>
<th>Steps</th>
<th>Variable</th>
<th>$B$</th>
<th>$SE$</th>
<th>$\beta$</th>
<th>CI</th>
<th>$R^2$</th>
<th>$F$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>IQ Score</td>
<td>.200*</td>
<td>.083</td>
<td>.210</td>
<td>.035</td>
<td>.365</td>
<td>.04</td>
<td>5.74</td>
</tr>
<tr>
<td></td>
<td>IQ Score</td>
<td>.201*</td>
<td>.084</td>
<td>.212</td>
<td>.036</td>
<td>.366</td>
<td></td>
<td>.04</td>
</tr>
<tr>
<td></td>
<td>$SD$ Attractiveness</td>
<td>.068</td>
<td>.076</td>
<td>.123</td>
<td>-.082</td>
<td>.218</td>
<td></td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>$M$ Attractiveness</td>
<td>.039</td>
<td>.037</td>
<td>.119</td>
<td>-.035</td>
<td>.112</td>
<td>.08</td>
<td>2.09</td>
</tr>
<tr>
<td></td>
<td>$SD$ Intelligence</td>
<td>.011</td>
<td>.084</td>
<td>.018</td>
<td>-.155</td>
<td>.178</td>
<td></td>
<td>.036</td>
</tr>
<tr>
<td></td>
<td>$M$ Intelligence</td>
<td>-.066</td>
<td>.047</td>
<td>-.158</td>
<td>-.159</td>
<td>.027</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .001. **p < .01. *p < .05. Two-tailed probabilities.

Note: Dependent variable is the attractiveness-intelligence halo (correlation) calculated for each participant.

---

⁴ A standard Pearson’s $r$ correlation was also analysed for the ‘average halo’ metric, even though it was not normally distributed. The pattern of findings did not change; that is, no significant correlations or non-significant trends were observed between any of the variables mentioned.
5.7 Discussion

As predicted, perceived intelligence and perceived attractiveness of facial stimuli were strongly correlated. In a more significant contribution to the literature, our findings suggest that participants varied in the degree of endorsement of the perceived intelligence-attractiveness relationship. Specifically, participants who scored higher on intelligence tests were more likely to endorse the perceived intelligence-attractiveness relationship. No significant interaction between individual intelligence and scale use (SD and M ratings for perceived intelligence and attractiveness) was found, and the relationship between participants’ intelligence score and their perceived intelligence-attractiveness halo remained significant after controlling for differences in participant scale use. We found no relationship between gender or age of participant and the intelligence-attractiveness halo.

We also found that there was no significant relationship between own intelligence and the “average halo.” The difference between the “individual halo” and “average halo” is salient; the “individual halo” is calculated by correlating individual ratings of intelligence and individual ratings of attractiveness, whereas the “average halo” correlation is calculated based on a correlation between an individual’s ratings of intelligence and the average attractiveness (of all 126 participants in the sample) of stimuli. Indeed, literature has suggested that individual face preferences are unique to each individual and are primarily related to individual differences in learning what is attractive from their environment (Germine et al., 2015), but may also be related to differences in culture (Penton-Voak et al., 2004), self-similarity resemblance (Alvergne, Faurie, & Raymond, 2009; Debruine, 2002), or self-evaluation (Kandrik & DeBruine, 2012; Little et al., 2001). Thus, measuring how similar an individual’s intelligence ratings are to average attractiveness ratings is more of a measure of conformity to what is considered by most as attractive, rather than a measure of how similar each person’s perceptions of intelligence and attractiveness are. Alternatively, one could interpret the positive relationship between intelligence and the “individual halo” and null relationship with the “average halo” to suggest that intelligent people are more likely to use heuristics like the anchoring effect. If so, one would expect intelligent individuals to be more likely to give similar ratings for other pairs of attributes.
Our findings concur with prior research showing that ratings of children’s intelligence from facial images correlate strongly with their rated attractiveness (Asch, 1946; Dion et al., 1972; Langlois et al., 2000). This is a troubling finding given the potential implications in education, as research suggests expectations of children’s intelligence can influence their academic performance up to five years after measurement of classroom teachers’ expectations (de Boer et al., 2010). Expectancy effects can be particularly robust in groups of students who are considered to be at risk of poor academic attainment (Hinnant, O’Brien, & Ghazarian, 2009; Sorhagen, 2013). Hence, our findings that individuals with higher intelligence have a stronger endorsement of the intelligence-attractiveness halo may be particularly worrying if we assume that individuals in university education and with varying levels of cognitive performance (such as those studied here) will progress to various employments including teaching.

Our findings highlight one particular individual difference that is related to the endorsement of the intelligence-attractiveness relationship. It is possible that people who score higher on intelligence tests are either more susceptible to a halo effect (perhaps due to the anchoring effect) or find faces they believe to be intelligent more attractive (perhaps due to a bias for self-similarity). We suggest that future experimental designs should use more controlled and concentrated experiments to examine these and other possible explanations. For instance, the intelligence test used in this study (ICAR) was very short; a more comprehensive intelligence assessment may shed light on which cognitive abilities are most strongly related to the perceived intelligence-attractiveness halo, thus narrowing potential explanations. Future research could also benefit from comparing participants’ intelligence-attractiveness halo to participants’ use of anchoring (see Tversky & Kahneman, 1974 for a classic study on measuring anchoring), heuristics, and stereotypes. Conversely, research could investigate whether people who are rated by others as more attractive are more likely to endorse the perceived intelligence-attractiveness correlation (halo).

Regardless of rationale (kin selection or anchoring effect) or directionality (i.e., whether perceived intelligence influences perceptions of attractiveness or vice versa), our findings reveal that individuals differ in their vulnerability to this bias. Further studies on other individual differences potentially related to the endorsement of the perceived intelligence-attractiveness halo are necessary, especially in the context of expectations of
children who are arguably most vulnerable to perceptions of intelligence, given the consequences of expectancy effects discussed. Davis (1951) said "the eye sees only what the mind is prepared to comprehend," thus, we should strive to better understand the root of biases in an effort to see past them.
Chapter 6: Blinded by Beauty: Attractiveness Bias and Accurate Perceptions of Academic Performance

This chapter is based on research that is under peer-review in an academic journal:

6. Abstract

Despite the old adage to not judge a book by its cover, facial cues often guide first impressions and these first impressions guide our decisions. Literature suggests there are valid facial cues that assist us in assessing someone’s health or intelligence, but such cues are overshadowed by an attractiveness halo whereby desirable attributions are preferentially ascribed to attractive people. The impact of the attractiveness halo effect on perceptions of academic performance in the classroom is concerning as this has shown to influence students’ future performance. We investigated the limiting effects of the attractiveness halo on perceptions of actual academic performance in the faces of 100 university students. Given the ambiguity and various perspectives on the definition of intelligence and the growing consensus on the importance of conscientiousness over intelligence in predicting actual academic performance, we also investigated whether perceived conscientiousness was a more accurate predictor of academic performance than perceived intelligence. Perceived conscientiousness was found to be a better predictor of actual academic performance when compared to perceived intelligence and perceived academic performance, and accuracy was improved when controlling for the influence of attractiveness on judgments. These findings emphasize the misleading effect of attractiveness on the accuracy of first impressions of competence, which can have serious consequences in areas such as education and hiring. The findings also have implications for future research investigating impression accuracy based on facial stimuli.
6.1 Introduction

A review by Langlois et al. (2000) suggested that people regularly make judgments based on appearance and argued that if humans were not biased to judge others on their appearance, they would not need to remind their children not to judge books by their covers (p. 408). While frequently warned against judging a book by its cover, the field of face perception is filled with evidence that suggests that the face contains a substantial amount of information for evaluators to infer traits. For instance, Kramer and Ward (2010) found that four of the Big Five personality traits, as well as physical health, were perceived with some limited accuracy from internal facial features alone and three of the Big Five traits were accurately perceived (just above chance) from just one side of the face. Similarly, Penton-Voak, Pound, Little, and Perrett (2006) found that there was some limited accuracy in perceptions of extraversion, emotional stability, and openness to experience when presented with images of composite faces (combining the faces of people with the same personality). Also, research by Little, Burt, Penton-Voak and Perrett (2001) found that evaluators were differentially attracted to faces depending on personality traits desired in a partner; that is, if a trait is desired then faces perceived to possess that trait are found more attractive than faces which do not possess that trait (p. 1107). Such research highlights potential accuracy in face perception and the relationship between perceived personality traits and attractiveness.

Indeed, when investigating the accuracy of perceived intelligence (Kleisner et al., 2014) and of perceived health (Kalick, Zebrowitz, Langlois, & Johnson, 1998) in faces, it was found that accuracy was improved to a level above chance when controlling for attractiveness bias. The attractiveness halo effect in which desired personality traits are ascribed to attractive people over unattractive people (Dion et al., 1972) seems to influence the use of attractiveness as a cue when attempting to accurately perceive health or intelligence in faces and is in turn, limiting people’s accuracy. The relationship seems to reflect a suppression effect, in which the suppressor (perceived attractiveness) is correlated with the other predictor variable (perceived health or intelligence), but is not related to the dependent variable (actual health or intelligence), so when this noise (relationship between attractiveness and perceived health or intelligence) is controlled for, the accuracy in perceptions of actual health or intelligence is increased (Conger, 1974; Maassen & Bakker,

### 6.1.1 Actual intelligence and attractiveness

Kleisner, Chvatalova, and Flegr (2014) reported accurate perceptions of intelligence in men’s but not women’s faces. It is important to note that a significant relationship between perceived and actual intelligence was only evident after statistically controlling for perceived attractiveness, though perceived attractiveness itself was not found to be a valid cue to actual intelligence. Kleisner et al. (2014) argue that one of the reasons accurate estimations of intelligence are demonstrated in men but not women may be due to the stronger effect of the attractiveness halo in perceptions of female intelligence. These findings highlight the pervasive and detrimental influence of attractiveness on accuracy in attributions.

For decades, researchers have debated the accuracy in perceived intelligence and whether attractiveness is a valid cue to actual intelligence (Denny, 2008; Jackson, Hunter, & Hodge, 1995; Kanazawa, 2011; Kleisner et al., 2014; Langlois et al., 2000). A study by Zebrowitz, Hall, Murphy, and Rhodes (2002) found that judgments of intelligence from faces were more accurate than chance for images from childhood, puberty, and middle adulthood, but not more accurate than chance in adolescence or late adulthood. Zebrowitz et al. (Zebrowitz et al., 2002) discussed how facial attractiveness might relate to actual intelligence based on various potential paths: (a) biological, with good genes being inherited; (b) environmental, including the impact of nutrition and healthcare; (c) influence of intelligence on grooming and health decisions; (d) and a self-fulfilling prophecy, in which attractive people are expected to be smarter and given greater opportunities to become smarter. A later study by Zebrowitz and Rhodes (2004) investigated the relationship between facial attractiveness and actual intelligence in the upper and lower halves of the attractiveness distribution and reported that, consistent with the “bad genes hypothesis,” facial attractiveness was a valid cue to actual intelligence only in the lower half of the attractiveness distribution. As discussed in Chapter 3, consistent with the anomalous face overgeneralization hypothesis, attractiveness was used to guide impressions of intelligence across the entire attractiveness distribution (Zebrowitz & Rhodes, 2004). Thus, participants were accurate in judging intelligence based on
attractiveness, but only because faces perceived as unattractive were judged as having low intelligence. These findings are consistent with the "bad genes" hypothesis, which implies that faces perceived as very unusual or unattractive may be an indicator of poor genetic fitness.

A more recent study by Mitchem et al. (2014) highlights several problems in previous research investigating attractiveness and intelligence, namely publication bias (systematic publishing of results different from unpublished studies), inconsistencies in definitions of intelligence and attractiveness, research design flaws, and small sample sizes. They conducted research on the largest sample to date, utilizing a twin dataset and independently collected measures of facial attractiveness and general intelligence. They found no support for a relationship between actual intelligence and perceived facial attractiveness.

### 6.1.2 Attractiveness and academic performance

Research has also investigated the potential relationship between perceived attractiveness and *actual* academic performance, with no clear consensus. Some investigations have showed that students who are perceived as more attractive achieve higher grades and higher scores on standardized achievements tests (Felson, 1980; Salvia et al., 1977; Singer, 1964). Other studies failed to find any relationship (Clifford, 1975; Sparacino & Hansell, 1979).

Nonetheless, the relationship between perceived attractiveness and *perceptions* of academic performance is clear. A meta-analysis conducted by Dusek and Joseph (1983) scrutinized 14 studies investigating physical attractiveness and its relation to teacher expectancy. The review concluded that perceived facial attractiveness is significantly correlated with teacher expectations of academic performance and positive personality attributes. For example, a cornerstone study by Clifford and Walster (1973) indicated a significant correlation between physical appearance and teacher expectations. A similar study also suggested a positive correlation between teachers' ratings of attractiveness and expectations of children's skills (Kenealy, Frude, & Shaw, 2001), showing that teachers judged children rated as more attractive as more social, confident, popular, academically strong, and more likely to become leaders than students who were rated as less attractive.
Another meta-analytic review by Ritts, Patterson, and Tubbs (1992) found that students perceived as attractive are more likely than students perceived as unattractive to be ascribed positive educational traits. Specifically, students perceived as attractive were judged as more intelligent, having more academic potential, and having better grades. It was also noted that other variables such as gender, race, and knowledge of past performance also influenced expectations, but were not significant moderators to the attractiveness influence (Ritts et al., 1992). Consequently, while there is little consensus and weak supporting evidence for a relationship between perceived attractiveness and \textit{actual} intelligence or academic performance, there is convincing research documenting the relationship between perceived attractiveness and \textit{perceived} intelligence and academic performance.

\textbf{6.1.3 Accuracy in face perception}

Research suggests extroversion can be accurately perceived after only a 50 millisecond exposure to a face (Borkenau, Brecke, Möttig, & Paelecke, 2009), strength can be accurately estimated from faces independent of height, weight, and age (Sell et al., 2009), and the dark triad of personality (Machiavellianism, narcissism, and psychopathy) can be accurately perceived in composites of expression-neutral facial images (Holtzman, 2011). Note here that accuracy does not imply a large effect size; accuracy may be significant, but with performance only slightly above chance. Nonetheless, this limited accuracy is still somewhat impressive given the lack of conventional information (i.e., information about behavior) that we typically think affects such judgments; thus, the effects may be small, but they are still noteworthy. Todorov, Olivola, Dotsch, and Mendes-Siedlecki (2015) suggests that little time is needed to arrive at a consensus on social attributions from faces, yet many studies overstate the validity of these attributions. There are various perspectives on why and how such social attributions from faces are made that explain the potential both for accuracy and for limitations in accuracy (see Chapter 1).

Biological cues may shed light on how people are rating social judgments at above-chance accuracy from neutral-expression facial images alone. For instance, research suggests the shape of a face (sexual dimorphism, discussed in Chapter 1) is related to the current (Lefevre et al., 2013; Penton-Voak & Chen, 2004; Pound, Penton-Voak, & Surridge, 2009) and prenatal (Fink et al., 2005) levels of testosterone. Research has also
suggested that adiposity is closely associated with circulating testosterone (Rantala et al., 2013) and that adiposity has been shown to be related to perceived health and attractiveness, as well as measures of actual cardiovascular health and proneness to respiratory illness (Coetzee et al., 2009). Further, facial symmetry, and sex typicality in face shape have been shown to be related to disease resistance (Thornhill & Gangestad, 2006). Similarly, an average face shape may signal health, as abnormalities that make a face look slightly different from the average may be caused by genetic or environmental stress (Zebrowitz & Rhodes, 2004, p. 169). Carotenoid coloration in the face has also been found to signify quality of current diet (Stephen, Law Smith, et al., 2009). The face can also provide clues to recent sleep history, with those who are sleep deprived having less eyelid-openness and more downward mouth curvature than those that are well rested (see Chapter 4; Sundelin et al., 2013).

6.1.4 Health, attractiveness and over-generalization

Clearly, the face provides a variety of cues to hormones, health, and sleep status. One thing all of these cues have in common is their relationship to attractiveness. Namely, research investigating attractiveness and the "good genes" theory has argued that facial symmetry (Perrett et al., 1999), averageness (Langlois & Roggman, 1990), sexual typicality (Perrett et al., 1998), eyelid-openness and mouth-curvature (Axelsson et al., 2010), carotenoid coloration in the face (Stephen, Law Smith, et al., 2009), and adiposity (Coetzee et al., 2009) may be attractive because of their relationship to health (see Chapter 1; Barber, Arnott, Braithwaite, Andrew, & Huntingford, 2001; Gehrman et al., 2011; Rhodes, 2006).

The link between potential cues to health in the face and perceived attractiveness is one explanation for the "attractiveness halo effect". Research suggests this preference for attractive (or healthy looking) individuals appears early in infancy, with infants as young as two-months old gazing longer at attractive faces over unattractive or unusual looking faces (Slater et al., 1998, 2000). It is unknown whether or not such preferential looking reflects early learning (Bushnell, 2001; De Haan, Johnson, Maurer, & Perrett, 2001; Pascalis, de Schonen, Morton, Deruelle, & Fabre-Grenet, 1995). Further, Langlois, Roggman, and Reiser-Danner (1990) found that 12-month-old infants would play longer, have more involvement, experience less distress and withdrawal, and seem to exhibit more
pleasure when interacting with attractive people as compared to unattractive people. Also noteworthy is the degree of agreement regarding facial attractiveness. Specifically, studies have shown consistency between men and women regarding opinions of facial attractiveness (Rhodes, 2006). Surprisingly, agreement on facial attractiveness is apparent even across different countries (Langlois et al., 2000; Tovée et al., 2006).

In an attempt to investigate whether facial attractiveness provides evidence of actual health, which may partially explain this positive bias toward attractive people, Kalick, Zebrowitz, Langlois and Johnson (1998) found that evaluators' perceptions of attractiveness are actually poor predictors of current or future actual health. While attractive faces were mistakenly rated as healthier than their peers, the correlation between perceived health and actual health increased when attractiveness was statically controlled, implying that attractiveness suppresses the accurate recognition of health.

This improvement in accuracy of health judgments after controlling for attractiveness is similar to the improved accuracy of intelligence judgments when the attractiveness halo is statistically controlled (Kleisner et al., 2014). Indeed there is evidence to suggest a relationship between various health factors and cognitive or intellectual performance. Specifically, it has been found that phobic anxiety (Okereke & Grodstein, 2013), trait anxiety (Bishop, 2009; Eysenck et al., 2007), drug use (Sim, Simon, Domier, Richardson, & Rawson, 2001), diabetes (Gregg, 2000; Munshi et al., 2006), poor sleep (Ohayon & Vecchierini, 2005), and frailty (Yassuda et al., 2012) have been negatively associated with both health and cognitive function in older individuals. Similarly, exposure to chronic aircraft noise (Haines, Stansfeld, Job, Berglund, & Head, 2001), infection with parasitic worms (Jardim-Botelho et al., 2008) and food insufficiency (Alaimo, Olson, & Frongillo Jr, 2001) have been found to negatively impact health and cognitive performance in children. Given the close relationship between actual health, actual cognitive performance, and perceived attractiveness, facial cues to health might also be cues to both attractiveness and cognitive ability (see general "fitness factor" discussed in Chapter 1 and 3), leading to correlations between attractiveness and perceived competence. Such correlation might lead to overgeneralization and inaccurate perceptions of academic ability in healthy individuals based spuriously on attractiveness. Hence we explore whether or not the "blinded by beauty" phenomenon found in perceptions of health.
6.1.5 Theories of intelligence and academic performance

Given the controversy over definitions of intelligence and differences in theories of intelligence (Neisser et al., 1996) it is likely that, in addition to being limited by the attractiveness halo, accurate perceptions of intelligence are also limited by variation in understanding on the meaning of the term “intelligence.” While someone who agrees with a fixed theory of intelligence believes there is little a person can do to change their actual intelligence, someone with a growth theory of intelligence argues that intelligence can change over time with the appropriate environment (Blackwell et al., 2007; Dweck et al., 1995; Hong, Chiu, Dweck, Lin, & Wan, 1999).

Perceptions of academic performance from faces are likely to suffer similar inconsistencies in evaluator perspectives of what factors most influence academic performance. While research has consistently shown that intelligence predicts academic performance (Chamorro-Premuzic, 2007), it is well documented that the personality trait of conscientiousness is a stronger predictor of academic performance than intelligence (Chamorro-Premuzic & Furnham, 2004; Chamorro-Premuzic, 2007). Hence, it could be argued that asking evaluators to assess academic performance from faces would yield just as much ambiguity as attributions of intelligence, as consensus would be adversely affected both by disagreement in fixed vs. growth theories of intelligence, and by different perspectives on how much academic performance relies on intelligence versus conscientiousness (see Chapter 2).

Research on the Intelligence Competence Theory (ICT) further undermines consensus of perceived academic performance by suggesting that people who are less intelligent compensate by becoming more conscientious to reach their goals (Moutafi, Furnham, & Crump, 2003; Moutafi et al., 2004). Thus, some might think a person with a less intelligent looking face is more academically able because the person may work harder to get better grades. Previous studies have highlighted consensus and accuracy of perceptions of most of the Big Five personality traits from faces, yet conscientiousness is sometimes (Little & Perrett, 2007), but not always correctly detected (Kramer & Ward, 2011; Penton-Voak et al., 2006). Given the relationship between actual conscientiousness...
and academic performance (compared to intelligence), we explore whether perceptions of conscientiousness are more likely to predict actual academic performance than perceptions of intelligence.

6.2 Research Questions

Research investigating perceptions of academic performance has primarily been concerned with exploring the potential of attractiveness to be a valid predictor of academic performance (Clifford, 1975; Felson, 1980; Murphy, Nelson, & Cheap, 1981; Salvia et al., 1977; Singer, 1964; Sparacino & Hansell, 1979) and exploring the effects of perceived academic performance on students’ actual performance in the future (Clifford & Walster, 1973; Dusek & Joseph, 1983; Kenealy et al., 2001; Ritts et al., 1992). No research that we are aware of has investigated the potential accuracy of perceptions of actual academic performance from faces when controlling for the attractiveness halo. Given the different perspectives and theories of the term “intelligence” (Dweck et al., 1995; Dweck & Henderson, 1989; Neisser et al., 1996) and the varying perspectives on how much intelligence predicts academic performance compared to conscientiousness (Moutafi et al., 2003, 2004), we hypothesize that evaluators will be more accurate in perceiving actual academic performance when specifically asked to rate conscientiousness than when asked to rate the more ambiguous terms “intelligence” or “academic performance.”

Further, it is possible that attractiveness detracts from accuracy in perceptions of academic performance, much as attractiveness can detract from accuracy in perceptions of health and intelligence. While there are various seemingly logical explanations for why attractiveness could be a valid cue to academic performance, the empirical evidence for a link between the two is extremely weak and perhaps only existing in the lower half of the distribution (i.e., driven by potential outliers with genetic or developmental problems affecting both appearance and cognitive ability). We hypothesize an “attractiveness halo” in which attractiveness is not linked to actual academic performance but is significantly correlated with perceptions of academic performance. Further, we hypothesize that controlling for the misperceptions about attractiveness may improve accuracy in perceptions of academic performance.
We argue that this effect of controlling for attractiveness takes the form of a classic type of suppression (see Conger, 1974; Maassen & Bakker, 2001; MacKinnon et al., 2000; Mavor et al., 2009; Tzelgov & Henik, 1991). In classical suppression, the suppressor is unrelated to the variable of interest but is related to the predictor, and therefore the shared variance between the predictor (in this case, perceived conscientiousness, intelligence or academic performance) and the suppressor (attractiveness) is unrelated to the outcome measure (actual academic performance). By controlling for this irrelevant variance in the predictor, the strength of the association between the predictor and outcome variable increases. In other words, controlling for attractiveness may reveal a “blinded by beauty” phenomenon similar to that found in health (Kalick et al., 1998) and intelligence (Kleisner et al., 2014) judgments.

6.3 Method

6.3.1 Facial stimuli

Students from the University of St Andrews were recruited to take part in an experiment called “Influences in the Perception of Intelligence in Faces” as part of a larger data collection. One hundred of the most standardized (e.g., clean shaven, neutral expression and head posture) Caucasian faces between the ages of 18 and 24 ($M_{age} = 20.85$, $SD = 2.15$; 67 females, 33 males) were chosen as stimuli. The original image collection contained more women than men and removal of males with beards enhanced the gender bias. Nonetheless, we maximized the number of stimuli available for judgments to maintain power in the analysis. Selection of standardized faces was done blind to their academic performance. Todorov and Porter (2014) highlight significant differences in person impressions within multiple facial photos of the same person due to random variation and discusses how this can influence accuracy of personality inferences based on faces. Thus, it was important to select the most standardized stimuli. All of the stimuli photographs of participants used were taken under standardized lighting conditions and camera set-up; individuals had their hair pulled back, did not wear any kind of make-up or
jewelry, and were instructed to pose with a neutral facial expression. Face images were aligned on left and right pupils. Images were then resized and cropped (1608 x 2584 pixels) so that an equal proportion of hair and neck was exposed in each.

6.3.2 Academic performance measures

All participants consented to releasing their academic performance records for the purpose of this research. Academic records were accessed via the University's database. Academic performance at the University of St Andrews is marked on a 20-point scale reported to one decimal place for final module grades. An average academic performance was calculated by taking the Grade Point Average (GPA) across every year weighted by every module credit completed by the student. Participants varied in their course of study and the number of modules completed based on their year and semester of study (63 in Sciences, 37 in Arts; 44 first- and second-year undergraduates, 39 third- and fourth-year undergraduates, and 17 in postgraduate courses). Accordingly, methods of evaluation (e.g., exam, essay, and dissertation) varied.

---

1 Due to the findings in Chapter 4, which suggest subtle differences in “neutral” mouth curvature can influence perceptions of intelligence, we checked for any outliers in our stimuli with regards to mouth curvature. No outliers were found.
6.4 Participants

6.4.1 Face raters

Four separate groups of participants were recruited and paid via Amazon Mechanical Turk to obtain ratings of perceived attractiveness, intelligence, conscientiousness, and academic performance (no other face ratings were obtained for this study). Table 6.1 shows the demographics of each participant group. Differences in sample sizes were based on differences in the number of participants completing the task while the link was live on Amazon Mechanical Turk and the number of exclusions. Consistent with methods in previous chapters, participants who reported their ethnicity as different from “white Caucasian” were excluded when calculating the average ratings of perceived attractiveness, intelligence, conscientiousness, and academic performance, as stimuli presented were Caucasian and judgments of other ethnicities may be more susceptible to stereotypes (Zebrowitz et al., 2007). Analysis was re-run with all participants and there were no differences in the pattern of findings; that is, all significant results remain significant, and all non-significant results remain non-significant.

Table 6.1
Sample information

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>M Age</th>
<th>SD Age</th>
<th>Exclusions</th>
<th>Total Sample</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attractiveness</td>
<td>40.16</td>
<td>12.44</td>
<td>5</td>
<td>32</td>
<td>F = 11 M = 21</td>
</tr>
<tr>
<td>Intelligence</td>
<td>40.00</td>
<td>8.99</td>
<td>12</td>
<td>25</td>
<td>F = 16 M = 9</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>42.32</td>
<td>12.17</td>
<td>8</td>
<td>20</td>
<td>F = 10 M = 10</td>
</tr>
<tr>
<td>Academic Performance</td>
<td>38.28</td>
<td>12.29</td>
<td>16</td>
<td>47</td>
<td>F = 22 M = 25</td>
</tr>
</tbody>
</table>

Note: Each participant group reflects a separate group of raters for one perception task. Female is represented by F and male by M.

6.5 Procedures

6.5.1 Face rating task

Evaluators first previewed all stimuli with each image displayed for one second. The stimuli were then re-presented so that participants could rate the face on the focal trait
for each sample: perceived attractiveness, intelligence, conscientiousness, or academic performance. Faces were presented in random order. To ensure the paid participants were not quickly and hastily clicking through images, images were presented for at least one second before participants were allowed to continue to the next image, but no maximum response time was enforced. Evaluators then completed a questionnaire inquiring about their age, gender, and ethnicity.

Facial ratings were done on a 7-point scale with endpoints according to the face rating task: attractiveness endpoints were not at all attractive to very attractive; perceived intelligence endpoints were not at all intelligent to very intelligent; perceived conscientiousness endpoints were not at all conscientious to very conscientious; and perceived academic performance endpoints were very low academic performance to very high academic performance.

Participants who rated perceived academic performance were presented with a statement at the top of each facial image presented asking, “Please rate how well you think this person does in University compared to the other people presented.” Participants who rated perceived conscientiousness were presented with a statement at the top of each facial image presented that read “Conscientiousness is the personality trait of being thorough, careful, or vigilant — with the desire to do a task well. Based on the definition of conscientiousness provided — how conscientious do you perceive this face to be compared to the other faces presented?”

6.6 Results

6.6.1 Facial averages

Consistent with methods used in Chapter 4, facial averages of the faces presented were created to help the reader visualize perceptions of conscientiousness and the attractiveness halo. Facial averages (see Figure 6.1) were synthesized from the top 25% male and female faces (8 male and 16 female faces) and bottom 25% male and female faces with the highest and lowest scores on perceived conscientiousness (Benson & Perrett, 1992, 1993). These average images were then made symmetrical (see Perrett et al., 1999).
Figure 6.1. The images presented reflect the top and bottom 25% of faces perceived as most (left) and least (right) conscientious. The attractiveness halo would suggest that faces perceived as most conscientious (left) would be more attractive than the faces rated as least conscientious (right).
### 6.6.2 Correlation matrix

An average score of perceived attractiveness, intelligence, academic performance and conscientiousness was calculated for each of the 100 faces based on the average of all the evaluator ratings. Table 6.2 gives the zero-order correlations between ratings and actual academic performance and the demographic variables of age and gender. There was a significant correlation between older age and higher actual academic performance and female faces were perceived as more attractive (see Table 6.2). As predicted, there was no relationship between attractiveness and actual academic performance \((r = 0.03)\), but a strong positive correlation between attractiveness and perceived intelligence \((r = 0.81)\), attractiveness and perceived academic performance \((r = 0.74)\), and attractiveness and perceived conscientiousness \((r = 0.81)\).

Table 6.2  
Zero-order matrix

<table>
<thead>
<tr>
<th>Actual Academic Performance</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td></td>
<td>.282**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Sex</td>
<td></td>
<td>-.098</td>
<td>-.011**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3. Attractiveness</td>
<td></td>
<td>.027</td>
<td>.296**</td>
<td>-.296**</td>
<td>1</td>
</tr>
<tr>
<td>4. Intelligence</td>
<td></td>
<td>.072</td>
<td>.302**</td>
<td>-.206*</td>
<td>.807***</td>
</tr>
<tr>
<td>5. Conscientiousness</td>
<td></td>
<td>.175b</td>
<td>.313**</td>
<td>-.360***</td>
<td>.812***</td>
</tr>
<tr>
<td>6. Academic Performance</td>
<td></td>
<td>.124</td>
<td>.308***</td>
<td>-.150</td>
<td>.738***</td>
</tr>
</tbody>
</table>

**\(p < .01\), *\(p < .05\), "\(p < .1\). Two-tailed probabilities.

Note: Sex is coded female = 0, male = 1. Correlations are based on 100 faces.

Given the high correlations between rated attributes (perceived attractiveness, perceived conscientiousness, perceived intelligence and perceived academic performance), we wanted to ensure that any statistical controls were based on sufficiently reliable measures and discriminability valid constructs (consistent with methods used in Chapter 4). Cronbach's alphas were calculated for perceived attractiveness (32 ratings; \(\alpha = 0.94\)), intelligence (25 ratings; \(\alpha = 0.86\)), academic performance (20 ratings; \(\alpha = 0.73\)), and conscientiousness (47 ratings; \(\alpha = 0.91\)). After correcting for attenuation due to measurement error (Nunnally, 1978; Osborne, 2003), the relationships between
attractiveness and perceived intelligence ($r = 0.90$), between attractiveness and perceived academic performance ($r = 0.89$), and between attractiveness and perceived conscientiousness ($r = 0.88$) were all marginally higher but do not indicate redundancy.

Partial correlations were conducted in which the influence of age of face, sex of face, and perceived attractiveness were controlled for\(^2\). Partial correlations revealed (see Figure 6.2) a significant correlation between perceived conscientiousness and actual academic performance ($r = 0.22, p = 0.035$). The partial correlations reveal no relationship between actual academic performance and perceived academic performance ($r = 0.13, p = 0.191$) or perceived intelligence ($r = 0.06, p = 0.544$).

Figure 6.2. This bar graph shows the increased accuracy of the different perceived competence variables when controlling for perceived attractiveness. The same pattern emerges when controlling for the additional variables of sex and age of face.

\(^2\) Findings do not change when controlling for only attractiveness in the partial correlation. Nor do they change when controlling for the combination of attractiveness and age or the combination of attractiveness and sex of face.
6.7 Discussion

There are three main findings. First, there was no first-order relationship between perceptions of conscientiousness, academic performance or intelligence and actual academic performance. Second, when controlling for the expected influences that age, sex and perceived attractiveness have on perceptions of competence (perceived conscientiousness, academic performance and intelligence), then the relation between perceived competence and actual academic performance increased in strength. Third, perceived conscientiousness was the single best face perception predictor of actual academic performance (outperforming perceived intelligence and perceived academic performance), and again accuracy was significantly improved when controlling for the suppressor variable of attractiveness.

As we expected, the form of the relationship is one of classic suppression in which there is some factor (perceived attractiveness) that is correlated with perceptions of conscientiousness, but not correlated with actual academic performance (Conger, 1974; Maassen & Bakker, 2001; MacKinnon et al., 2000; Mavor et al., 2009; Tzelgov & Henik, 1991). When this factor is controlled, the relationship between perceived conscientiousness and actual academic performance is increased (see Figure 6.3). It should also be noted that, although some previous literature suggests weak correlations between attractiveness and cognitive performance measures (Zebrowitz et al., 2002), in our study perceived attractiveness was not a valid cue to actual academic performance. These results suggest that we are “blinded by beauty” in a way in which we would be more accurate in our perceptions of academic performance from faces if we were not influenced by the attractiveness halo effect.
Figure 6.3. Suppression effect. This figure shows the noise in perceived conscientiousness (the overlap between perceived attractiveness and perceived conscientiousness) and how by suppressing this noise results in an improved predictor of actual academic performance (greater overlap between the remaining perceived conscientiousness and actual academic performance).

Given the amount of research on higher expectations and desired educational traits being ascribed to attractive students over unattractive students, it is not surprising that faces that were rated as more intelligent, having better academic performance and being more conscientious were also rated as more attractive (see composite faces in Figure 6.1). As predicted, there were high correlations between perceptions of attractiveness and perceptions of intelligence, conscientiousness, and academic performance, likely reflecting the strength of the attractiveness halo, as well as the similarities among these perceived competence measures (Todorov, Said, Engell, & Oosterhof, 2008). While there is less evidence to suggest perceptions of intelligence and academic performance are unique
constructs, the possibility that perceived conscientiousness and perceived attractiveness are not distinguishable empirically is dealt with in two ways: face validity of the items for which evaluators were clearly rating conscientiousness or attractiveness (the measures were unambiguous to the evaluators); and we calculated inter-rater reliabilities for conscientiousness and attractiveness ratings and even after correcting for attenuation due to measurement error, the correlations between these variables remained distinct (i.e., they were imperfectly correlated). Taken together, these elements suggest that these measures can be treated here as distinct constructs, and that they are measured with sufficient reliability to be distinguished empirically in this study. The high correlations do create potential for interpretative difficulties in multiple regression, and under such circumstances we find it important to emphasize the role of suppression in their relationship in a way that reflects the traditional understanding of the attractiveness halo.

Findings suggest that accuracy in perceptions of academic performance also increases with the clarity and validity of the question proposed. When controlling for attractiveness, age and sex, perceptions of conscientiousness in faces yielded above chance accuracy in predicting academic performance, but accuracy in predicting actual academic performance did not reach levels of statistical significance with perceptions of intelligence or perceptions of academic performance. Given the high correlations between these perceived competence measures, it is difficult to say for certain whether perceptions of conscientiousness are unique in their capacity to predict actual academic performance over and above perceptions of intelligence or academic performance. Rather, it seems perceptions of conscientiousness predict actual academic performance because, in comparison, it may be the least ambiguous competence construct. As previously argued, it is likely that individual differences in theories and understandings of intelligence can lead, on average, to less accurate perceptions of intelligence in faces. Likewise, perceived academic performance is possibly confounded by a combination of the ambiguities in the term intelligence (fixed vs. malleable) and the limited consensus on how much intelligence (in relation to conscientiousness) is necessary for high academic performance; hence the limited accuracy of perceived academic performance compared to perceived conscientiousness in predicting actual academic performance.

The improved accuracy in perceived conscientiousness predicting actual academic performance over perceived intelligence is also consistent with research that suggests that
actual conscientiousness is a stronger predictor of academic performance than actual intelligence (Chamorro-Premuzic, 2007). Further, the Intelligence Compensation Theory (ICT) suggests that conscientiousness acts as a coping strategy for relatively less intelligent people. While evidence for ICT is limited, some studies have found significant negative correlations between fluid intelligence and conscientiousness (Moutafi et al., 2003, 2004). Other studies have found a significant negative correlation between crystalized intelligence and conscientiousness (Wood & Englert, 2009). Thus, our findings of perceived conscientiousness better predicting actual academic performance in faces than perceived intelligence is consistent with literature suggesting actual conscientiousness is a better predictor than intelligence in predicting actual academic performance. Nonetheless, given the high correlations amongst the perceived competence variables explored (perceived intelligence, perceived academic performance and perceived conscientiousness), we must be cautious in claiming that only perceived conscientiousness is related to actual academic performance; rather, we argue that the specificity in rating tasks and the influence of attractiveness bias are worth considering when exploring validity of judgments based on faces.

The increased accuracy of academic performance in faces after controlling for attractiveness has important implications. Indeed, Olivola and Todorov (2010) showed that judges overweight aspects of appearance and would be more accurate in judging personality if face perception was ignored. However, facial impressions have consistently been shown to influence our opinions as well as bias decisions in politics (Little, Roberts, Jones, & Burriss, 2007), leadership (Rule & Ambady, 2008), law (Zebrowitz & McDonald, 1991), parental expectations and punishments on children (Zebrowitz, Kendall-Tackett, & Fafel, 1991), military rank promotion (Mueller & Mazur, 1996), and teacher evaluations (Ambady & Rosenthal, 1993). Clearly, the power of first impressions is critical and has repeatedly been shown to influence our opinions about a person.

Furthermore, research has found that femininity is considered more attractive than masculinity (Perrett et al., 1998) and that females perform better academically and stay in education longer than males (Gurian & Stevens, 2004), which likely leads to females being ascribed more desired educational traits over men. It is also well documented that older students do better on intelligence tests (Cerella & Hale, 1994; Fry & Hale, 2000) and do better academically than younger students. Moreover, crystalized intelligence and
perceptions of wisdom have shown to increase linearly with age (Clayton & Birren, 1980; Horn & Cattell, 1967), which would influence impressions of competence in older students (hence the intentionally limited university age range for facial stimuli presented). Our research suggests that when controlling for biases of attractiveness, age and sex, independently or collectively, accuracy of perceived academic performance is significantly improved.

Perhaps one of the most alarming consequences of using insufficient information to guide first impressions is the expectancy effect in education. The classic Pygmalion study conducted by Rosenthal and Jacobson (1968) suggests that expectations alone are capable of influencing the target’s actual performance. Specifically, the Pygmalion study found that students who were arbitrarily assigned the label “bloomers” (i.e., anticipated to show future promise) eventually scored higher on future tests than other students, even though the students labelled as “bloomers” were a random sample and not any more intelligent than the other students in the class. More recent research on expectancy effects by Sorhagen (2013) found that teachers’ inaccurate expectations of students in first grade was associated with students’ academic performance in high-school and that students from lower-income families were especially influenced by this bias. Likewise, De Boer, Bosker, and Van Der Werf (2010) defined expectation bias as the difference between observed and predicted teacher expectation and found a significant relationship between teacher’s expectation bias of students’ performance and actual performance five years later. Hence, perceptions of conscientiousness, intelligence, and academic performance may play a vital role in the classroom environment and in the success of a child’s education.

Future research in face perception can benefit from noting the significant differences in perception accuracy based on different theories of intelligence or competence. Perhaps more importantly, given the well-documented effects of expectations of academic performance on actual academic performance, our findings help emphasize the biased effects of perceived attractiveness on expectations of academic performance. While it seems unlikely that another person’s attractiveness can be filtered out when attempting to accurately perceive academic performance, the mere knowledge of the negative influence attractiveness has on accuracy may encourage less-biased practice; for perhaps the best antidote to deter unconscious bias is to make conscious the possibility of bias.
Chapter 7: Conclusions
7.1 Summary of Results

Previous studies have noted the strong relationship between perceptions of intelligence and attractiveness. While the attractiveness halo has positive influences on a range of other social judgments (e.g., trustworthiness, leadership), the influence of attractiveness on perceptions of intelligence was of particular interest in this thesis because of the impact of expectations of intelligence on students’ actual academic performance (de Boer et al., 2010; Rosenthal & Jacobson, 1968). With special attention to the influence of attractiveness, this thesis explores cues to perceived intelligence, potentially accurate perceptions of competence, and individual susceptibility to the halo.

Chapter 4 introduces and validates new facial measurements to investigate malleable cues to perceptions of intelligence that are independent of the attractiveness halo. The cues of eyelid-openness and mouth curvature were specifically investigated given their relationship to sleep-deprivation and low mood, both of which have been associated with decreased cognitive ability (Ellis et al., 1984; Thomas et al., 2000). The findings described in Chapter 4 support the hypothesis that more eyelid-openness and upturned mouth curvature increase ratings of perceived intelligence independent of attractiveness. The eye and mouth shape influences on intelligence attribution were found for the faces of both adults and children. Both eyelid-openness and mouth curvature were confirmed to affect perceptions of intelligence using digital image manipulations in which only these shape cues were altered. Subtle differences in eyelid-openness and mouth curvature were also found to influence perceptions of intelligence within images of individuals before and after sleep restriction.

Chapter 5 extends research findings that perceived intelligence and attractiveness correlate by examining individual differences that may be related to the strength of this correlation. As predicted, our findings suggest that participants varied in the degree of endorsement of the perceived intelligence-attractiveness relationship, with participants who scored higher on intelligence tests being more likely to endorse the perceived intelligence-attractiveness halo. This influence of evaluator intelligence remained significant even when controlling for potential differences among individuals’ scale use (standard deviation and mean of ratings for perceived intelligence and attractiveness).
Chapter 6 investigates the potential for individuals to assess academic performance accurately from facial images. While there was no relationship between actual academic performance and any of the perceived competence measures (perceived conscientiousness, perceived academic performance or perceived intelligence), the relation between perceived competence and actual academic performance increased in strength and became significant when controlling for the expected influences that age, sex, and perceived attractiveness have on perceptions of competence. Further, although other perceived competence measures (perceived academic performance, perceived intelligence) were highly correlated with perceived conscientiousness, perceived conscientiousness was the best face perception predictor of actual academic performance (outperforming perceived intelligence and perceived academic performance when controlling for attractiveness). This is interesting in light of previous research that shows that actual conscientiousness is a stronger predictor of academic performance than actual intelligence (Chamorro-Premuzic & Arteche, 2008; Chamorro-Premuzic, 2006, 2007).

7.2 Contributions, Limitations, and Future Work

The findings presented in this thesis highlight the strong relationship between perceptions of attractiveness and intelligence, using novel methods. Chapter 4 investigates the facial cues to alertness and to mood on perception by objectively measuring the eyelid-openness and mouth curvature of facial images. Perceptual studies in which participants’ ratings of eyelid-openness and mouth curvature significantly correlated with the calculated measurements, thus validating the objective measurements of eyelid-openness and mouth curvature. The influence of eyelid-openness and mouth curvature on perceptions of attractiveness and intelligence highlights the importance of considering (and perhaps controlling for) the impact these cues may have on other social judgments. For instance, given the well-established influence of valence on trait attributions like trustworthiness (see Oosterhof & Todorov, 2008), it may be beneficial to control for any influence of mouth curvature when investigating the possibility that other cues like facial width to height measures (Stirrat & Perrett, 2010) or family resemblance (Debruine, 2002) affect trustworthiness impression. Hence, controlling for mouth shape is particularly important in cases where mouth curvature may systematically relate to the variable of interest. For
example, more dominant males may adopt a more down-turned mouth curvature during photography compared to less dominant males.

Similarly, an underlying assumption of most work investigating social judgments from facial images is that faces have a “neutral expression,” but by measuring subtle differences in mouth curvature and eyelid-openness, future research can more objectively ensure faces presented do not deviate significantly in regard to hints of positive or negative expression (as was done in Chapter 6). The use of measurements can also establish the relative importance of these shape cues when several cues can affect social judgments independently. For example, skin color, tiredness, demeanor, and level of adiposity may all contribute to judgments of apparent health.

As a further contribution, when investigating the role of mouth curvature and eyelid-openness on attributions, we directly manipulated these cues independently while keeping all other aspects of the face constant. While finding an effect of eyelid-openness and mouth curvature across natural images of adults and children and within the same individual after sleep-restriction was significant, the potential for these cues to be working in combination with each other or other cues was possible. Indeed, the direct manipulation of each cue independently allowed for stronger inferences about each cue than analysis of natural images alone.

Another key contribution stemming from Chapter 4’s finding is that the cues of eyelid-openness and mouth curvature influence perceptions of intelligence independent of attractiveness. We ensured eyelid-openness and mouth curvature were not simply influencing perceived intelligence through their impact on attractiveness by conducting mediation analysis. Future research may consider using mediation analysis to examine the direct and indirect effects of particular facial cues on social judgments. For example, facial symmetry is known to influence perceptions of attractiveness and also be a significant predictor of actual intelligence (Banks et al., 2010; Miller, 2000; Penke et al., 2009); researchers might use mediation analysis to investigate whether facial symmetry is a cue to perceived intelligence independent of attractiveness.

Clearly, the influence of attractiveness on social judgments is robust. While much research (in addition to the research presented in this thesis) has established that people perceive faces they consider intelligent as attractive (and vice versa), Chapter 5
investigates this phenomenon at the individual level by creating a metric based on the
correlation between a participant’s ratings of attractiveness and ratings of perceived
intelligence for a set of children’s faces. While we used children’s faces as stimuli because
of the well-established influence of perceptions of children’s intelligence on expectations
and future academic performance, further research should examine the viability of one’s
own intelligence influencing the endorsement of the perceived intelligence-attractiveness
relationship in adults.

Additionally, while our findings suggest that individuals who scored higher on
intelligence tests had a stronger endorsement of the perceived intelligence-attractiveness
halo (or correlation), future research may benefit from investigating other individual
differences that may influence endorsement of the intelligence-attractiveness halo, like
own-attractiveness. Likewise, researchers might use a similar method to investigate the
basis of age- or gender-biased stereotypes more generally. By measuring the association or
correlation between face, gender or age and a social judgment (such as employability or
suitability for leadership), one can determine the degree of bias at an individual level. As
investigated in Chapter 5, further analysis can then be made of the extent to which a
particular trait (own age, gender, education) is associated with sexism or ageism.

While attractiveness strongly relates to perceptions of intelligence, research
suggests attractiveness and actual intelligence have no actual relationship (Mitchem et al.,
2014). Indeed, this lack of validity makes it the attractiveness halo. Nonetheless, Chapter
6 finds that perceptions of competence correlated with actual academic performance, once
we statistically controlled for the attractiveness halo, and that perceived conscientiousness
correlated more strongly than perceived intelligence or perceived academic performance
itself. These findings have important implications in relation to potentially accurate trait
attributions from faces and helps emphasize the limiting effects of overgeneralizations.
Future research may consider investigating whether perceptions of personality from facial
stimuli alone would be more accurate if desired personality traits (agreeableness,
conscientiousness, etc.) were not influenced by the attractiveness halo.

It is important to note that statistically controlling for an overgeneralization like
attractiveness does not imply that a person can actually control for their bias and be more
accurate, but rather emphasizes the point that overgeneralizations are typically blinding.
Future research is necessary to explore whether individuals can be less biased if they attempt to control for their own overgeneralizations. Indeed, literature suggests increasing self-control strength (through exercises like using one’s non-dominant hand or refraining from cursing) can reduce the effect of suppressing stereotypes (Gailliot, Plant, Butz, & Baumeister, 2007), yet such self-regulation takes effort (Gordijn, Hindriks, Koomen, Dijksterhuis, & Van Knippenberg, 2004; Richeson & Shelton, 2003; Richeson & Trawalter, 2005). Further, while all of the empirical studies in this thesis have potential implications in areas such as education and employment, the experiments presented are analogue studies that do not sample teachers or employers and as such, future research is needed to investigate these samples specifically.

Additionally, Chapter 6’s findings emphasize the significant differences among word usage in perceptual face rating tasks and accuracy of perceptions. While perceived conscientiousness, academic performance, and intelligence are related to perceived competence (and indeed are all highly correlated), specificity among these tasks resulted in increased accuracy. Collectively, research investigating judgment validity from facial images could benefit from examining specific perceptions (e.g., perceived blood pressure versus perceived health in assessing actual blood pressure) and controlling for well-known overgeneralizations to emphasize their limiting effects.

### 7.3 Final Conclusions

This thesis investigates the relationship between perceived intelligence and attractiveness with particular interest in implications in education, hiring, and other social judgments that are shaped by expectations and first impressions. The findings from Chapter 4 suggest that regardless of attractiveness, an individual can alter his or her perceived intelligence by changing eyelid-openness (by getting more sleep) and mouth curvature (by avoiding a subtle frown). Similarly, evidence from Chapter 5 suggests that individuals with higher intelligence may be more prone to an attractiveness halo effect, while Chapter 6 emphasizes the limiting effects of these types of overgeneralizations by showing that controlling for the attractiveness halo reveals better assessments of competence. In a next step, research could investigate these effects in a more natural
environment (such as the classroom or workplace) and investigate whether awareness of the findings presented can alter behavior and perceptions for the better.
References


Buss, D. M. (1985). Human mate selection: Opposites are sometimes said to attract, but in fact we are likely to marry someone who is similar to us in almost every variable. *American Scientist, 73*(1), 47–51. doi:10.2307/27853061


Byrne, D., London, O., & Reeves, K. (1968). The effects of physical attractiveness, sex, and attitude similarity on interpersonal attraction. *Journal of Personality, 36*(2), 259–271.


Cerella, J., & Hale, S. (1994). The rise and fall in information-processing rates over the


Cunningham, M. R. (1986). Measuring the physical in physical attractiveness. Quasi-


Denny, K. (2008). Beauty and intelligence may or may not be related. *Intelligence, 36*(6), 616–618.


in the perception of age, attractiveness, and health in female faces. *Evolution and Human Behavior, 27*(6), 433–442. doi:10.1016/j.evolhumbehav.2006.08.007


control of human body height, weight and BMI in five ethnically. Annual Human
Genetics, 62, 307–322.

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

Gordijn, E. H., Hindriks, I., Koomen, W., Dijksterhuis, A., & Van Knippenberg, A.
(2004). Consequences of stereotype suppression and internal suppression
30(2), 212–24. doi:10.1177/0146167203259935


Gottfredson, L. S., & Deary, I. J. (2004). Intelligence longevity, predicts but and why
Current Directions in Psychological Science, 13(1), 1–4.

doi:10.1017/S1464793102006085

Grammer, K., & Thornhill, R. (1994). Human (Homo sapiens) facial attractiveness and
sexual selection: The role of symmetry and averageness. Journal of Comparative

decline among older women? Archives of Internal Medicine, 160(2), 174–180.
doi:10.1001/archinte.160.2.174

Association, 40(1), 85–90. doi:10.1037/h0033731

doi:10.1037/h0046827

Guilford, J. P. (1977). Way beyond the IQ: Guide to improving intelligence and


257–302.

aircraft noise exposure, stress responses, mental health and cognitive performance
doi:10.1017/S0033291701003282

Theoretical Biology, 7(1), 1–16. doi:10.1016/0022-5193(64)90038-4


doi:10.1037/0012-1649.31.3.464


doi:10.1016/j.jaad.2007.07.040


Short, L. A., Mondloch, C. J., McCormick, C. M., Carré, J. M., Ma, R., Fu, G., & Lee,


Evolution and Human Behavior, 27(6), 443–456. doi:10.1016/j.evolhumbehav.2006.05.004


Appendix 1: ICAR Sample Intelligence Test
Cognitive Ability Test (ICAR)

Permission to include a shortened ICAR sample test in this thesis was granted via personal communication with the ICAR team, but for test security, only four example items are presented (one item from each item type category) and no answers to the items are provided. For more information about the psychometric properties of this 16-item ICAR Sample Test, please refer to The international cognitive ability resource: Development and initial validation of a public-domain measure (Condon & Revelle, 2014). These items are examples of the four item types described in Chapter 2 and Chapter 5: verbal reasoning, letter and number series, matrix reasoning, and mental rotation. Test questions were presented to participants one-by-one in a random order.

**Verbal Reasoning:**

1. (VR) Zach is taller than Matt and Richard is shorter than Zach. Which of the following statements would be most accurate?

   Richard is taller than Matt, Richard is shorter than Matt, Richard is as tall as Matt, It's impossible to tell, Richard is taller than Zach, Zach is shorter than Matt, Zach is shorter than Matt, none of these, I don't know

**Letter and Number Series:**

2. (LN) In the following alphanumeric series, what letter comes next?
   V, Q, M, J, H...

   E, F, G, H, I, J, none of these, I don't know
Matrix Reasoning Items

3. (MR) Please identify which of six geometric shapes presented as response choices will best complete the stimuli.

A, B, C, D, E, F, none of these, I don't know

Three-Dimensional Rotation Items

4. (R3D) Please identify which of the response choices is a possible rotation of the target stimuli.

A, B, C, D, E, F, G, H, none of these, I don't know
Appendix 2: Ethics Application Forms
19 February 2014

Ethics Reference No: PS10828

Project Title: Influences in the perception of intelligence in faces – Online Study

Researcher's Name: Sean Talaman

Supervisor: Professor David Perrett

Thank you for submitting your application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 12th February 2014. The following documents were reviewed:

1. Ethical Application Form 18/02/2014
2. Advertisement 18/02/2014
3. Participant Information Sheet 18/02/2014
4. Consent Form 18/02/2014
5. Debriefing Form 18/02/2014
6. Questionnaire 18/02/2014
7. Data Management Plan 18/02/2014

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the "Guidelines for Ethical Research Practice" https://www.st-andrews.ac.uk/utrec/guidelines/ are adhered to.

Yours sincerely

Convener of the School Ethics Committee

[Signature]

Prof D. Perrett (Supervisor)
School Ethics Committee

School of Psychology & Neuroscience, St Mary's Quad, South Street, St Andrews, Fife KY16 9J9
Email: psychethics@st-andrews.ac.uk Tel: 01334 462071

The University of St Andrews is a charity registered in Scotland: No SC013532
Project Title | Influences in the perception of intelligence in faces – Online Study
---|---
Researcher’s Name | Sean Talamas
Supervisor | Professor David Perrett
Department/Unit | School of Psychology & Neuroscience
Ethical Approval Code (Approval allocated to Original Application) | IPS10828
Original Application Approval Date | 18 February 2014
Amendment Application Approval Date | 06 April 2015

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 24th March 2015. The following documents were reviewed:

1. Ethical Amendment Application Form | 06/04/2015
2. Participant Information Sheet | 06/04/2015
3. Debriefing Form | 06/04/2015
4. Questionnaire | 06/04/2015

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension. You will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/medsci/UTREC/guidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convener of the School Ethics Committee

Cts | Prof D. Perrett (Supervisor)
School Ethics Committee

School of Psychology & Neuroscience, St Mary's Quad, South Street, St Andrews, Fife KY16 9JP
Email | psychology@st-andrews.ac.uk Tel 01334 462071

The University of St Andrews is a charity registered in Scotland No SC013532
<table>
<thead>
<tr>
<th>Project Title</th>
<th>Influences in the perception of intelligence in faces – Online study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher's Name</td>
<td>Sean Talamanas</td>
</tr>
<tr>
<td>Supervisor</td>
<td>Dave Perrett</td>
</tr>
<tr>
<td>Department/Unit</td>
<td>School of Psychology &amp; Neuroscience</td>
</tr>
<tr>
<td>Ethical Approval Code</td>
<td>PS10828</td>
</tr>
<tr>
<td>Original Application Approval Date</td>
<td>18 February 2014</td>
</tr>
<tr>
<td>Amendment Application Approval Date</td>
<td>23 April 2014</td>
</tr>
</tbody>
</table>

**Ethical Amendment Approval**

Thank you for submitting your amendment application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 23rd April 2014. The following documents were reviewed:

1. Ethical Amendment Application Form 23/04/2014
2. Participant Information Sheet 23/04/2014
3. Consent Form 23/04/2014
4. Questionnaire 23/04/2014

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the "Guidelines for Ethical Research Practice" (http://www.st-andrews.ac.uk/medic/UTREC/guidelines%20.files%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Cc Prof D. Perrett (Supervisor)
School Ethics Committee
Project Title: Influences in the perception of intelligence in faces – In lab study

Researcher's Name: Sean Talmanas
Supervisor: Professor David Perrett
Department/Unit: School of Psychology & Neuroscience

Ethical Approval Code: TP10867
Original Application Approval Date: 04 March 2014
Amendment Application Approval: 09 June 2014

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 28th May 2015. The following documents were reviewed:

1. Ethical Amendment Application Form: 09/06/2014
2. Advertisement: 09/06/2014

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/muhs/UTREC/guidelines%20for%20research%20practice.pdf) are adhered to.

Yours sincerely,

Convener of the School Ethics Committee

Ces Prof D. Perrett (Supervisor)
School Ethics Committee
Ethical Amendment Approval

Thank you for submitting your amendment application which was considered by the Psychology & Neuroscience School Ethics Committee on the 6th November 2014. The following documents were reviewed:

1. Ethical Amendment Application Form 06/11/2014
2. Participant Information Sheet 06/11/2014
3. Consent Form 06/11/2014
4. Debriefing Form 06/11/2014

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three-year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension of you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the ‘Guidelines for Ethical Research Practice’ (http://www.st-andrews.ac.uk/media/1/UTREC/guidelines%20for%20research%20practice.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ces  Prof D. Perrett (Supervisor)
School Ethics Committee

School of Psychology & Neuroscience, St Mary's Quad, South Street, St Andrews, Fife KY16 9JP
Email: psyethics@st-andrews.ac.uk Tel: (0) 334 4632731

The University of St Andrews is a charity registered in Scotland No SC013552
Ethical Amendment Approval

Thank you for submitting your amendment application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 12th March 2014. The following documents were reviewed:

1. Ethical Amendment Application Form 12/03/2014
2. Questionnaire 12/03/2014

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTREC/guidelines/0204feb52008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Cc Prof D. Perrett (Supervisor)
School Ethics Committee
Ethical Amendment Approval

Thank you for submitting your amendment application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 23rd April 2014. The following documents were reviewed:

1. Ethical Amendment Application Form 23/04/2014
2. Advertisement 23/04/2014
3. Participant Information Sheet 23/04/2014
4. Consent Form 23/04/2014

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical considerations, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the ‘Guidelines for Ethical Research Practice’ (http://www.st-andrews.ac.uk/media/UTRECguidelines%28Feb%29/08.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Cox Prof D. Perrett (Supervisor)
School Ethics Committee
### Ethical Approval

Thank you for submitting your amendment application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 27th January 2015. The following documents were reviewed:

1. Ethical Amendment Application Form 28/01/2015

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the ‘Guidelines for Ethical Research Practice’ ([http://www.st-andrews.ac.uk/medic/UTREC/guidelines%20for%20research%20practice%20%282006%29.pdf](http://www.st-andrews.ac.uk/medic/UTREC/guidelines%20for%20research%20practice%20%282006%29.pdf)) are adhered to.

Yours sincerely

Convener of the School Ethics Committee

Ces Prof D. Perrett (Supervisor)
School Ethics Committee

---

School of Psychology & Neuroscience, St Mary's Quadrangle, South Street, St Andrews, Fife KY16 9JH
Email: psyEthics@st-andrews.ac.uk Tel: 01334 462071
The University of St Andrews is a charity registered in Scotland: No SC013532

---

<table>
<thead>
<tr>
<th>Project Title</th>
<th>Individual Differences in the Perception of Facial Attractiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher's Name</td>
<td>Sean Tzamandas</td>
</tr>
<tr>
<td>Supervisor</td>
<td>Professor David Perrett</td>
</tr>
<tr>
<td>Department/Unit</td>
<td>School of Psychology &amp; Neuroscience</td>
</tr>
<tr>
<td>Ethical Approval Code</td>
<td>P11165</td>
</tr>
<tr>
<td>Original Application Approval Date</td>
<td>30 September 2014</td>
</tr>
<tr>
<td>Amendment Application Approval</td>
<td>28 January 2015</td>
</tr>
</tbody>
</table>
University of St Andrews
University Teaching and Research Ethics Committee
Sub-committee

2 October 2014

<table>
<thead>
<tr>
<th>Ethics Reference No:</th>
<th>PS1165</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please quote this ref on all correspondence</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Project Title:</th>
<th>Individual Differences in the Perceptions of Facial Attractiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researchers' Names:</td>
<td>Sean Talman, Hugo Trad, Louise Milligan, Emma Martin, Rebecca Munns, Olivia Ives and Katie Harrison</td>
</tr>
<tr>
<td>Supervisor:</td>
<td>Professor David Perrett</td>
</tr>
</tbody>
</table>

Thank you for submitting your application which was considered at the Psychology & Neuroscience School Ethics Committee meetings held on the 2nd July and 24th September 2014. The following documents were reviewed:

1. Ethical Application Form 30/09/2014
2. Advertisement 30/09/2014
3. Participant Information Sheet 30/09/2014
4. Consent Form 30/09/2014
5. Debriefing Forms 30/09/2014
6. Questionnaires 30/09/2014
7. Data Management Plan 30/09/2014

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' [https://www.st-andrews.ac.uk/utrec/guidelines] are adhered to.

Yours sincerely,

Convener of the School Ethics Committee

Coo School Ethics Committee
Prof D Perrett (Supervisor)

School of Psychology & Neuroscience, St Mary's Quad, South Street, St Andrews, Fife KY16 9LP
Email: psychethics@st-andrews.ac.uk, Tel: 01334 462071

The University of St Andrews is a charity registered in Scotland: No SC001352