The impact of social learning on first impressions across development

Adam Thomas Eggleston

PhD

University of York

Psychology

December 2021

Abstract

First impressions play a prominent role in guiding human decision-making. Impressions from faces, in particular, demonstrate the ability to influence voting decisions, criminal sentencing outcomes and leadership opportunities. Given the wide-ranging outcomes of first impressions, a significant amount of research has attempted to investigate their origins and development. One of the most prominent views in the field is that first impressions are governed by an innate mechanism, able to account for the early development, automaticity and speed of first impressions. The aim of my doctoral studies was to explore alternative accounts of the origins of first impressions. Specifically, across three empirical chapters I investigate the potential role of social learning on the formation and maintenance of first impressions. The results of the empirical work in Chapter 2 highlighted the plausibility of social learning as one route to first impressions. Results showed that impressions from culturally learned cues share characteristics of impressions from physical cues in their automaticity, speed and early development (by 6 years of age). Chapters 3 & 4 went on to explore how consistent first impressions may emerge early in development. Chapter 3 demonstrates the ease with which children, from at least 5 years of age, can use the non-verbal behaviour of others to infer traits such as niceness. Traits which were then transferred to a similar looking, but novel, target. Chapter 4 highlights the potential for parent-child conversations to facilitate the formation of face-trait mappings. Both parents and children made trait inferences when discussing faces, without explicit instruction to do so, with parents actively reinforcing their children’s face-trait mappings. The potential role of social learning in the formation and maintenance of first impressions is discussed as well as possible avenues for future research.

List of Contents

List of Figures

Acknowledgements

I would like to extend my deepest gratitude to Dr Harriet Over. As a supervisor I could not have asked for anyone more intelligent, encouraging, and generous with their time. Conducting research under her supervision has been a genuine joy and an incredible learning experience.

I am also grateful to have worked closely with so many amazing people in the Social Origins Lab (past and present). As a lab they are incredibly caring and talented. I cannot imagine how I would have managed without their support.

Special thanks go to those who I have had the privilege to collaborate with, Dr Richard Cook, Dr Cade McCall, Dr Jon Flavell, Dr Steven Tipper and Dr Elena Geangu. All your insights and guidance have helped me to grow as an academic over the past three years.

A big thank you also goes to my parents and my sister. My sister is the reason I started studying psychology and has been a constant source of support over the past three years. My parents have always been on hand to lend and ear and their unconditional love and support have helped me more than they can ever know.

Finally, I would like to thank my incredible wife. I truly could not have completed this thesis without her. She has read every word I have written at least five times, with minimal complaints, and continues to be a never-ending well of positivity and encouragement.

Author’s Declaration

I, Adam Eggleston, declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

The empirical work presented in this thesis has been published in the following peer-reviewed journals:

Eggleston, A., Flavell, J. C., Tipper, S. P., Cook, R., & Over, H. (2021). Culturally learned first impressions occur rapidly and automatically and emerge early in development. *Developmental Science*, *24*(2), e13021.

A. Eggleston, H. Over and R. Cook developed the study concept. All authors contributed to the study design. D. Ma granted the first author use of their stimuli. Testing and data collection was performed by A. Eggleston and J.C. Flavell. Data analysis and interpretation were performed by A. Eggleston and H. Over. Second coding was completed by B. Fisher. All authors contributed to and approved the final version of the manuscript for submission.

Eggleston, A., Geangu, E., Tipper, S. P., Cook, R., & Over, H. (2021). Young children learn first impressions of faces through social referencing. *Scientific Reports*, *11*(1), 1-8.

A. Eggleston, H. Over and R. Cook. All authors contributed to the study design. K.A. Dalrymple granted the first author use of their stimuli. Testing and data collection was performed by A.E. Data analysis and interpretation were performed by A. Eggleston, H. Over. Second coding was completed by B. Fisher. All authors contributed to and approved the final version of the manuscript for submission.

Eggleston, A., McCall, C., Cook, R., & Over, H. (2021). Parents reinforce the formation of first impressions in conversation with their children. *Plos one*, *16*(8), e0256118.

A. Eggleston, H. Over and R. Cook developed the study concept and contributed to the study design. D. Lundqvist and A. Todorov granted the first author use of their stimuli. Testing and data collection was performed by A. Eggleston. Data analysis and interpretation were performed by A. Eggleston, H. Over and C. McCall. Reliability coding was completed by L. Tunstall and S. Allen. All authors contributed to and approved the final version of the manuscript for submission.

Primary Supervisor Statement

I am listed as a co-author on the three empirical papers which make-up the main body of this thesis.

In each of the reported studies, the work is primarily that of Mr. Adam Eggleston. For each paper, Adam compiled the relevant stimuli, completed all of the data collection and coded and analysed the summary data. Adam wrote the first draft of each paper.

Harriet Over (10/12/2021)

Chapter 1: General Thesis Introduction

Humans are adept at categorising others based on minimal information, effortlessly categorising others based on their age, gender and race (Karnadewi & Lipp, 2011). Perhaps more surprisingly we also categorise people on their personality traits based on their physical facial characteristics, (Oosterhof & Todorov, 2008). These first impressions exert a powerful influence over behaviour and once made can be difficult to override (Cikara & van Bavel, 2014; Todorov et al., 2005). Adults who appear untrustworthy are less likely to be offered positions in interview settings (Olivola et al., 2014a) and are more likely to face harsh sentences in criminal justice situations (Wilson & Rule, 2015). Adults who look competent are more likely to be elected to public office (Ballew & Todorov, 2007). Importantly, whilst some of these first impressions demonstrate above chance levels of accuracy (Bonnefon et al., 2017), many others bear little resemblance to the actual personalities of the individuals being judged (Rule et al., 2013). Given the implications of these wide-spread and potentially misleading impressions it is imperative that research focuses on understanding their development and psychological origins.

The objective of my doctoral studies was to explore the development of first impressions, with specific focus on the mechanisms able to explain their origins. The work presented in Chapters 2-4 aims to contribute to the work on first impressions by offering a close examination of competing theories that consider their roots. Additionally, I wish to explore specific routes through which first impressions could be acquired and disseminated within a culture. The overall goal of this work is to clarify the developmental origins of first impressions to help inform future interventions that could potentially mitigate their harmful effects.

First Impressions from Faces

A Brief History of First Impressions

Research into first impressions from faces has a long history. Some of the earliest literature we have that theorises a link between a person's face and character is accredited to the ancient Greeks. In Physiognomica, a treatise attributed to Aristotle, parallels are drawn between a person's character and their resemblance to certain animals, an example of this being that a man who resembles a lion is brave (Evans, 1969). Whilst later works abandoned this animal analogy the science of Physiognomy endured, experiencing a significant revitalisation in the 18th century. At this point in history physiognomy had been given rules and was believed to be a skill one could possess and hone (Shortland, 1986). Still developing as a discipline, the 19th century saw the rise of a more scientific practice, seeing scientists such as Francis Galton utilising new methodologies like composite pictures to try to categorise criminals (Galton, 2007). The influence of this early work is evident in modern day research whether it be the use of face morphing and averaging techniques or the focus on inferences of criminality.

It is clear that historical research has informed modern day techniques, but it has also helped to reveal the consequences of first impressions. A powerful example of this is the use of physiognomy in Germany during WWII. At this time the Nazi Party were trying to achieve racial purity and physiognomy granted them a "scientific" basis for the spread of racist propaganda (Gray, 2004). Galton's work and other physiognomic texts were compulsory reading in schools during this time, used to try to create physical racial stereotypes that existed to feed growing xenophobia. For example, in Galton's work Jewish males are described as having a "cold, scanning gaze" (Collins, 1999). This example, though extreme, reveals the consequences of believing first impressions to be accurate and diagnostic, as well as demonstrating how our impressions can be manipulated through propaganda (Hansen, 2009).

Dimensions of First Impressions

Since this early work, some research has focused on identifying the core dimensions that make up our first impressions and so underlie these historical consequences. Early evidence from Oosterhof & Todorov (2008) revealed, using data-driven computational modelling, a two-dimensional model of first impressions. To arrive at this two-dimensional model, unconstrained descriptions of faces were collected from participants. From these descriptions a subset of 14 trait dimensions were identified and participants were asked to rate the same faces based solely on these attributes. A principal components analysis found that two dimensions were sufficient to account for a large proportion of the variability in trait judgements: valence and dominance. The valence dimension was characterised as a component whereby positive judgements (e.g. Caring) had positive loadings and negative judgements (e.g. Unhappy) had negative loadings. The highest loadings for the second component came from judgements of dominance, aggressiveness and confidence and was interpreted as a dimension for dominance evaluations. This two-factor model of first impressions is thought to represent an evolutionary preparedness to judge a person's threat level and so promote survival (Collova et al., 2019; Oosterhof & Todorov, 2008).

Since this original study the two-dimensional structure of first impressions has been replicated (Oh et al., 2019). However, the results from these models can vary by their input and so work has also been done to expand the input into the model. When participants gave judgements on 100 traits a four-dimensional structure of first impressions emerged: warmth, competence, youthfulness and femininity (Lin et al., 2021), mirroring the results from other work (Sutherland et al., 2013). Using the same data-driven techniques as Oosterhof and Todorov (2008), but with a wider age range, and more diverse or “ambient” stimuli, Sutherland et al’s (2013) findings supported the original two-dimensional model. However, similar to Lin and colleagues (2020), the use of ambient images also resulted in the finding of a possible third dimension labelled youthful/attractiveness. This dimension has been interpreted as having evolutionary origins in sexual selection (Sutherland et al., 2013). A key difference between this third dimension is that it is based on observable physical features in contrast to trustworthiness and dominance which are more abstract personality features that must be inferred. As this thesis is concerned with these unobservable aspects of a person’s character, attention will remain on this two-dimensional model.

One outstanding question based off the literature reviewed so far is whether this two-dimensional model can be replicated across cultures. To address this gap Sutherland et al.(2018), using the same data driven techniques as Oosterhof & Todorov (2008), built models of both Chinese and British perceivers’ impressions of own and other race faces. A striking similarity was found between the dimensions used by both cultures in their impression formation. It was found that both cultures structure their judgements around approachability, though judgements of Chinese participants were less clearly structured around capability than were the judgements of British participants (Sutherland et al., 2018). The new models from British and Chinese participants also resulted in a dimension more akin to general capability/experience rather than the dimension of dominance reported in earlier work (Oosterhof & Todorov, 2008; Sutherland et al., 2013). This early work led to the conclusion that first impressions are facilitated by a set of universally used dimensions.

To test the generalisability of models of first impressions further, an impressive large scale cross-cultural replication of Oosterhof & Todorov's (2008) original two-dimensional model has recently been undertaken (Jones et al., 2021). This project saw participants (N = 11, 481) from eleven world regions, featuring 41 countries, rate a more diverse set of faces on the same thirteen traits used in the original research. Results found that, depending on the world region, the number of reduced dimensions ran between two and four (Jones & Kramer, 2021; Jones et al., 2021). Some have reported that finding variability in the number of dimensions used across regions challenges the universality of the two-dimensional model (Xie et al., 2021). Others argue it is evidence that the original two-dimensional structure generalises well across world regions, especially given that the principal dimension in all regions was that of valence/trustworthiness (Todorov & Oh, 2021).

Overall, work on the dimensions used to form first impressions repeatedly report valence/trustworthiness and power/dominance as key factors through which others are mapped and therefore judged (Lin et al., 2021; Oh et al., 2019; Oosterhof & Todorov, 2008; Sutherland et al., 2013). This indicates that there may be a common structure on to which traits are mapped. An important note however is that this does not mean that people use these dimensions in the same way. Research has shown how an individual's conceptual beliefs about the relationship between different traits influences their first impressions and that this can differ across individuals and contexts (Oh et al., 2019; Stolier et al., 2020; Stolier et al., 2018;

Sutherland et al., 2020a). Additionally, recent work has demonstrated that the content of judgements made may interact with the type of face being judged. For example, when judging children's faces the main dimension used to form judgements is still akin to valence (niceness) but unlike other work the second dimension to emerge is shyness (Collova et al., 2019).

Contemporary Research on the Consequences of First Impressions.

Literature on first impressions is starting to reveal that the consequences of evaluating others across dimensions such as trustworthiness and dominance are not just relegated to history. This section aims to review the wealth of research detailing the modern-day repercussions of first impressions. Specifically, this section will focus on work following three central themes: crime & punishment, election outcomes and leadership & success. Work in these areas demonstrates how wide-spread consequences of first impressions are. This work highlights the importance of research investigating how first impressions can be manipulated, for example through social media or use of cultural cues (Chapter 2).

Crime and Punishment

As already described, early work on first impressions often focused on trying to identify supposedly undesirable individuals from their facial characteristics. This is an enduring theme, with research frequently focusing on consensus in participant judgements of a target's "criminality" (Flowe, 2012; Funk et al., 2016). These judgements of criminality have been shown to remain stable after repeated exposures and after minimal exposure of just 100ms (Klatt et al., 2016). However, in real cases both judge and jury are exposed to more information than a defendant's face, having access to detailed information about the defendant and the case. When participants are privy to the details of a target's crime it interacts with facial characteristics in interesting ways. For example, a "baby-faced" adult is more likely to be found guilty of an offense resulting from negligent actions, but less likely to be found guilty of an offense from intentional criminal behaviour than a mature-faced defendant (Berry & Zebrowitz-McArthur, 1988; Zebrowitz & McDonald, 1991).

One theory that considers how first impressions and evidence interact is the dangerous decisions theory. According to this theory, the evaluation of character traits from a defendant's face may influence "the interpretation and assimilation of incoming evidence" (Porter & ten Brinke, 2009). It is this bias in decision making, from first impressions, that research is beginning to demonstrate. Participants seem to require less evidence and report higher confidence when delivering a guilty verdict for untrustworthy verses trustworthy looking defendants (Porter et al., 2010). Going beyond the verdict, other work has shown the role first impressions continue to play in criminal sentencing. Reviewing real-world examples in the sentencing of convicted murderers, researchers from the US found that perceptions of untrustworthiness predicted the likelihood that a convict would receive a death sentence over a life sentence (Wilson & Rule, 2015). These studies depict a criminal-justice system where extremely fast judgements of a defendant's character traits have potential real-world consequences on the identification, judgement and sentencing of a defendant, even in extreme cases such as the death penalty.

Election Outcomes

The influence of first impressions on election outcomes is perhaps one of the most researched areas within the literature and is arguably the topic that sparked a resurgence of interest in the field. Early work by Todorov and colleagues (2005) had participants choose between two black and white photographs of winning and runner up candidates from US senate and house races. Participants were asked to judge in each case who they thought to be more competent and it was found that this judgement saw participants correctly predict the election outcome in 71.6% of Senate and 66.8% of House races (Todorov et al., 2005). This effect remained significant even when the viewing time of the faces was limited to 100ms (Olivola & Todorov, 2010; Todorov et al., 2005) and is a phenomenon not just confined to US politics. Studies using the same, or very similar, procedures have replicated these findings for elections in: Spain (Brusattin, 2012), Australia (Martin, 1978), Bulgaria (Sussman et al., 2013) and in newer democracies such as Mexico and Brazil (Lawson et al., 2010). Taking the results of these studies in to account it may seem that political parties would fare best if they simply put their most competent looking candidate up for election. However, outside an experimental setting, voters are afforded the opportunity to learn about specific candidates' policies and beliefs and candidates are not matched by gender, race or age. Therefore, to truly understand the consequences of first impressions on election outcomes contextual factors need to be taken into consideration.

One factor which may affect how strong the influence of first impressions are on election outcomes is the experience and knowledge of the voter. Work sampling Chinese rural constituencies, whose main source of information about a candidate is visual, found that they seem to rely on inferences of competence when making voting decisions (Wong & Zeng, 2017). In contrast, it was found that election outcomes in urban constituencies, where access to substantive information about candidates is more easily accessible, perceived competence of candidates was not found to be relevant (Wong & Zeng, 2017). The importance of the source of voter's political knowledge has also been demonstrated in western cultures. A US study, testing the relationship between appearance-based inferences and media exposure revealed that the influence of first impressions on voting patterns was highest amongst high TV/low political knowledge participants (Lenz & Lawson, 2011).

The power of media sources, such as TV, to influence our inferences of a target's character traits has been hypothesised as an important factor in forming and maintaining first impressions (Over & Cook, 2018). In the context of election outcomes this media influence could have powerful consequences. The media depictions of a target that a person sees may have a real impact on both their impression of that target and their voting behaviour and recent work goes some way in confirming this. A review of depictions of political candidates from online sources has shown that more favourable (warm, competent, happy) images of candidates can be found in line with the media source's political leanings (Hehman et al., 2012; Peng, 2018). Whether this bias is intentional is yet to be explored, however this research along with the other work reviewed in this section reflect a world where elections are very much influenced by first impressions and in which these impressions can be defined, and possibly manipulated, by the media a person consumes.

Leadership & Success

Career success in many contexts may be subjective, just as the success of a political candidate would be difficult to objectively measure. Because of this, research in the areas of leadership and success have focused on careers with clear, measurable hierarchies. A good illustration of this is military rankings, a career with a clear chain of command and where promotion is thought to be heavily dependent on ability. A pair of studies investigating the relationship between perceived dominance of West Point alumni and their career success found that dominant looking cadets were more likely to be promoted than their submissive looking peers (Mazur et al., 1984; Mueller & Mazur, 1996). Similarly, and more recently, Linke and colleagues (2016) found that within corporate hierarchies the highest positions are more likely to be taken by those perceived as more trustworthy. From this pair of studies, it seems that certain facial features may afford some a distinct personal advantage when it comes to career success.

Researchers have found evidence that first impressions cannot only predict personal success but also company success. Judgements of CEO leadership and power were shown to be positively correlated with multiple measures of firm success, such as company profits (Rule & Ambady, 2008a, 2009). An important question to ask is whether; (A) physical facial features are an accurate indicator of performance, or (B) physical facial features skew our perception of a person's ability/performance. A pair of studies looking directly at this question seem to support hypothesis B. First, by demonstrating that the "look" of the CEO does not accurately predict performance (Stoker et al., 2016). Secondly, by identifying that decisions on salary are influenced by perceived competence in external but not internal hiring decisions, where first impressions would be less important given the interviewers advanced knowledge of the candidate (Canace et al., 2020; Graham et al., 2017). Much of the work reviewed so far has focused on how first impressions effect hiring decisions at the level of upper management, where a person may be seen as the face of the company. However, biases from first impressions are not reserved for higher up positions, effecting general judgements of employability (Menegatti et al.,

2021) by those with and without hiring experience (Bjornsdottir & Rule, 2017; Filkuková & Jørgensen, 2020).

Whilst it does seem that physical facial features are an important factor in hiring and promotion, it does not seem to be the case that there is a "face of success" that would give a person an advantage across multiple domains. Numerous studies have demonstrated that different organisations value different perceived traits and that people are sensitive enough to these differences to be able to tell leaders from different organisations apart, e.g. business leader vs sports leader (Olivola et al., 2014a; Re & Rule, 2016, 2017). Furthermore, it seems that these valued traits may be adaptable even within an organisation, dependent on the situation (Little, 2014; Spisak et al., 2012). For example, participants were shown to prefer a masculine looking leader in a war-time scenario but a more feminine face in a peace-time scenario (Little, 2014).

Summary

Overall, the work reviewed here on the consequences of first impressions paints a world where major decisions are partially governed by snap judgements of character from physical facial characteristics. It also seems that these consequences are not unique to a single culture, with the literature demonstrating facial biases across world regions. Whilst prevalent, the impact of first impressions does seem to be partially mitigated by a person's knowledge of a target individual (Canace et al., 2020; Graham et al., 2017), suggesting room for interventions to combat their consequences. The wide-reaching effects of first impressions highlight it as an important area of research. If inaccurate, then first impressions have a deleterious effect on society, resulting in wrongful convictions, poor leadership choices and biased employment decisions. It is therefore imperative that the mechanisms behind first impressions are clarified to better understand how and when they are formed.

Accuracy

As discussed in the previous section, the accuracy of first impressions plays a key role in their function and the motivation for the research behind them. High accuracy would mean that first impressions are unproblematic and that the ability to make trait inferences is in fact an invaluable social tool. However, inaccurate trait inferences would mean that one of the key tools used to drive decisions making is unreliable, resulting in biased and potentially harmful behaviour.

Within the accuracy literature there is a substantial amount of work reporting above chance levels of accuracy in first impressions (Lin et al., 2018; Little et al., 2013; Porter et al., 2008; Slepian & Ames, 2016), including in judgements of children's "niceness" (Collova et al., 2020a). Aside from trustworthiness, research has also reported above chance levels of accuracy in judgements of a target's sexuality (Rule et al., 2009; Rule & Ambady, 2008b), political orientation (Olivola et al., 2012; Rule & Ambady, 2010) and level of prejudice (Alaei & Rule, 2019). Research mostly shows that accuracy is just above chance level (Bonnefon et al., 2015; Bonnefon et al., 2017), with some evidence that this level may increase following longer exposure to target individuals (Bonnefon et al., 2017; Carney et al., 2007) and with age (de Neys et al., 2015).

However, research on the accuracy of first impressions is currently fractured, with some researchers reporting that claims of accuracy are overstated and that researchers and policy makers should aim to reduce the influence of face-based judgements or "face-ism" (Olivola et al., 2014b). Work to support this has shown that for traits indicating trustworthiness, arguably the most important dimension on which faces are mapped (Todorov & Oh, 2021), there is no convincing evidence of accuracy (Dilger et al., 2017; Efferson & Vogt, 2013; Lavan et al., 2021; Todorov & Porter, 2014). Importantly, it seems that the level of accuracy may be domain specific, with researchers finding above chance accuracy for judgements of intelligence and extraversion but not for trustworthiness (Rule et al., 2013). The low levels of accuracy frequently demonstrated may not be a problem if people are aware of the times their judgements are more or less accurate (meta-accuracy). People would then be able to control their behaviour, only acting on the impression when they are highly confident it is correct. However, researchers have reported that participants confidence in their judgements and their accuracy are seemingly unrelated (Ames et al., 2010; Hong et al., 2021; Jaeger et al., 2020a).

Some other theories have focused less on the origins of first impressions as a whole and more on why studies are often finding any level of accuracy. One interesting theory, dubbed the Dorian Gray effect (Zebrowitz et al., 1998) is that personality may influence appearance. Some evidence for this comes from research indicating that your emotional disposition can be imprinted on your face over time (Adams et al., 2016; Malatesta et al., 1987). For example, people with a bad temperament may tense certain facial muscles associated with frowning more than those with a good temperament. Over time frequently used facial muscles are strengthened, resulting in visible physical differences between the two personality types. A second theory, which has recently garnered more support within the literature, is that to some extent accurate judgements are the result of a self-fulfilling prophecy (Hong et al., 2021; Li et al., 2017). For example, researchers confirmed that if a person was perceived as selfish they were treated more selfishly and, because of this treatment, behaved more selfishly (Haselhuhn et al., 2013).

Overall, research on the accuracy of first impressions suffers from various problems that could explain the lack of consistency in results. These include, contradictory results between different measures of the same trait (Little et al., 2013), high within-person variability when judging two pictures of the same person (Todorov & Porter, 2014) and biases resulting from stimuli selection (Todorov et al., 2015). With these issues in mind, it is hard to use the accuracy (or lack thereof) of first impressions to make inferences about their origins. One conclusion shared by many researchers is that although there may be truth in at least some first impressions (Bonnefon et al., 2017), the level of accuracy and meta-accuracy is not large enough for people to base important decisions on first impressions.

Theories of First Impressions

Evolutionary Theory

Evolutionary theories of first impressions are historically the most prominent amongst the literature and state that trait inferences are the product of an innately specified mechanism. This mechanism manifested as a distinct evolutionary advantage for our ancestors in the form of heightened threat detection and sexual selection (Schaller, 2008; Sutherland et al., 2018; Zebrowitz & Zhang, 2009). As stated in a previous section, a majority of trait inferences are reported as falling along three principal dimensions and researchers have attempted to highlight how sensitivity along these dimensions would constitute an evolutionary advantage. For example: trustworthiness, is used to assess whether a target’s intentions are good or bad; dominance, is used to assess the target’s ability to carry out their good or bad intentions; and youthful/attractiveness, is used to identify healthy sexual partners (Oosterhof & Todorov, 2008; Sutherland et al., 2013). However, it could be argued that judgements of attractiveness are fundamentally different than those of trustworthiness and dominance. The latter being categorised as personality traits offering cues to potential behaviours and the former categorised as a physical trait serving as a cue to overall health and reproductive fitness. Research has also highlighted this dichotomy, reporting how judgements of attractiveness may even precede trait judgements (Gutiérrez-García et al., 2018).

Evolutionary theories of first impressions are tasked with addressing why many trait inferences show only minimal levels of accuracy (Collova, et al., 2020a; Efferson & Vogt, 2013; Rule et al., 2013). It seems unlikely that detection of trustworthiness and dominance offered our ancestors a significant evolutionary advantage if our inferences are frequently inaccurate. However, some have argued that impressions merely have to be “minimally diagnostic” to be advantageous (Schaller, 2008), with psychological mechanisms following the smoke detector principle whereby failure to detect a threat is more dangerous than detecting a threat that is not present (Nesse, 2006). Therefore, the low accuracy in trait inferences could be an important function of the evolutionary mechanism, designed to be oversensitive in helping us avoid threatening social situations.

Alternatively, the overgeneralisation hypothesis accounts for the low accuracy of trait inferences by arguing that some trait inferences are the consequence of the overgeneralisation of adaptive mechanisms (Todorov, 2008; Zebrowitz et al., 2003). For example, an adaptive mechanism may be sensitive to a salient physical cue, such as smiling, used to judge a person’s trustworthiness. This mechanism however may also then be sensitive to subtle resemblances to this emotional expression. As a consequence of overgeneralisation, a person presenting a neutral expression merely resembling a happy expression will be judged more trustworthy than a person whose neutral expression resembles disgust or fear (Said et al., 2009, 2011). In line with the emotion overgeneralisation hypothesis, studies have repeatedly demonstrated how subtle emotional expression modulates trait judgements of neutral faces (Albohn et al., 2019; Albohn & Adams, 2021; Said et al., 2009; Zebrowitz et al., 2010). Aside from emotional-face overgeneralisation researchers have suggested, and provided evidence for, a wide array of other generalisations (baby-face, unfit-face, familiar-face) to explain a host of inaccurate judgements (Zebrowitz, 2017). In the case of baby-face overgeneralisation it has been found that targets whose features resemble those often associated with babies (e.g. large eyes, small nose) are more likely to be associated with traits such as naivety and honesty (Berry & McArthur, 1986). In unfit-face overgeneralisation, targets with asymmetrical or non-average features are more likely to be considered unattractive and unintelligent (Zebrowitz et al., 2003). Finally, familiar-face overgeneralisation posits that the evolutionary advantageous adaptation to distinguish friend from foe is overgeneralised so that even strangers who merely resemble familiar face are judged more positively (Zebrowitz et al., 2007).

Cultural Learning Theory

Recent work on first impressions has focused on the influence of top-down factors, such as individual experiences, beliefs and culture, on first impressions (Stolier, Hehman, & Freeman, 2018; Stolier, Hehman, Keller, et al., 2018)**.** Cultural learning models are able to account for these factors and so have gained more attention in recent years (Lee et al., 2021; Over et al., 2020). For example, the Trait Inference Mapping Framework (TIM), contends that first impressions are the result of learned associations between facial appearance and personality traits, associations which are acquired throughout ontogeny (Over & Cook, 2018). According to TIM one route through which face-trait mappings may be acquired is via direct experience and conscious evaluation, whereby repeatedly observing a person with a particular facial characteristic and displaying a trait related behaviour (e.g., performing well on a school test) causes an association to form in the brain between corresponding locations in a hypothesised “face space” and “trait space” (See Figure 1). Once acquired a mapping can then be automatically activated when encountering a stranger who physically resembles a previously encountered individual. This may result in a trait inference in the absence of any diagnostic behaviour. Together these methods for trait-face mappings form a dual-route model whereby associations can be formed either through explicit reasoning or automatically.

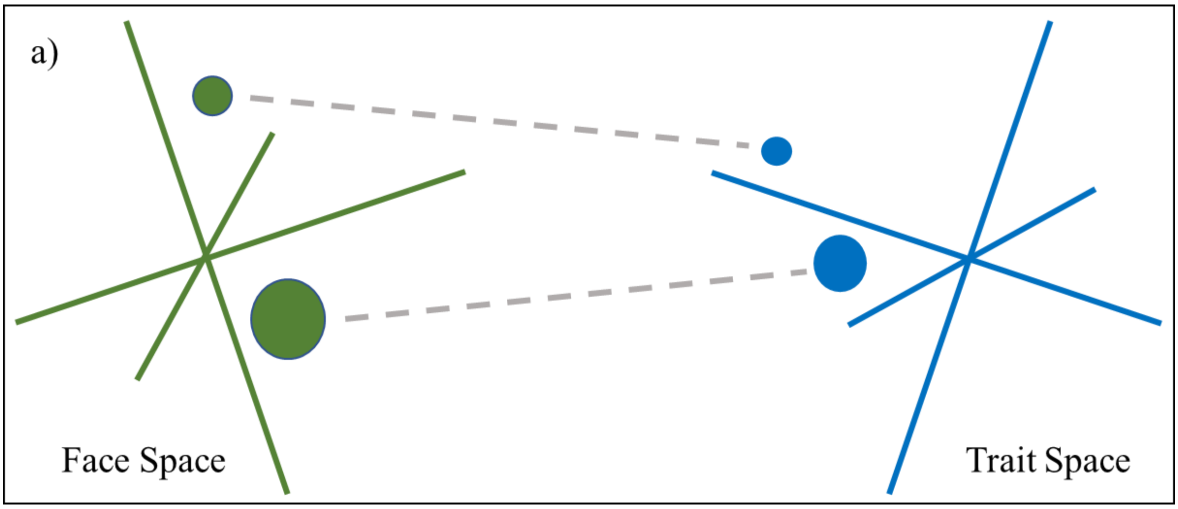

Figure 1: The Trait Inference Mapping Model (TIM) of first impressions (Over & Cook, 2018). According to TIM, first impressions are the product of learned mappings that allow excitation to spread from perceptual representations of face shape (points in face space) to representations of trait profiles that others may possess (points in trait space).

A possible criticism of learning models is that they seem to rely on veridical relationships between traits and behaviours, as face-trait mappings may only occur where a predictive relationship exists. A major strength of the TIM framework is that it can explain this paradox via cultural learning and innate physiological responses. First, according to Over and Cook (2018), various cultural devices (TV, fairy tales, books, propaganda) expose us to consistent face-trait mappings. For example, in animated movies protagonists are often presented with flawless skin whereas antagonists are more likely to be presented with some sort of dermatological condition (Ryan et al., 2018). Cultural learning, via media or face to face interactions, is one powerful route through which information can be passed down through generations (Nielsen et al., 2014, 2021). Face-trait mappings may propagate across generations via this method, either through explicit warnings and descriptions or through unconscious language use and non-verbal behaviours (Weisbuch et al., 2009).

Secondly, TIM proposes a small role for innate mechanisms within its framework that can help explain the consistency paradox. It has been documented that certain facial characteristics, as well as other stimuli, provoke predictable physiological responses. Attractive faces elicit sexual arousal and, in the same vain, spiders elicit a fear response (Rakison & Derringer, 2008; Rhodes, 2006). The physiological responses to certain stimuli are said to canalise particular face-trait mappings (Over & Cook, 2018), for example an attractive face may elicit positive feelings which then promote a positive interpretation of ambiguous behaviours. As a result, face mappings for attractive faces may more likely be associated with positively valenced traits. Recent work has begun to explore this relationship between physiological states and their influence on face-trait mappings. For example, when inducing stress in participants it was found that the perceived facial trustworthiness of neutral faces decreased (Toet et al., 2017). Other work has shown that when pairing faces with unrelated images that vary in valence participants are more likely to trust a composite image of the faces previously paired with high valence images compared to low valence (Kocsor & Bereczkei, 2017). Importantly facial affect may also elicit this type of stimulus-response, altering a person's physiological state and so guiding face-trait mappings. The evidence for this comes from a study that presented participants with a target neutral face and, using continuous flash suppression, an unseen face which was either smiling or frowning. Results showed that the unseen face influenced trait judgements of trustworthiness and likeability, regardless of whether or not the unseen face was the same identity as the target (Anderson et al., 2012). Together this work supports the concept that a mechanism, responsible for changes in physiological state to certain stimuli, may help to canalise the development of particular face-trait mappings. Furthermore, given that innate implies some sort of genetic basis, similar face-trait mappings in response to certain stimuli (e.g., smiling, frowning) would be expected across cultures, helping to address the consistency paradox within the TIM framework.

Summary

When considering cultural learning theories and evolutionary theories it is important to note that researchers genuinely agree that neither are independent of the other (Schaller, 2008). For example, as previously stated, TIM posits a small but important role of innate mechanisms (Over & Cook, 2018). Likewise, evolutionary theorists accept that first impressions are malleable and so influenced by learning (Ewing et al., 2019). It is however important to tease apart these theories to consider how predetermined first impressions are; or as (Hassin & Trope, 2000) put, “are we trapped in our faces?”. Given the consequences of first impressions, evolutionary theories may consider that early interventions to mitigate possible harmful effects of first impressions would be minimally effective. Cultural learning theories, however, would hypothesise the opposite.

Characteristics of First Impressions

In order to evaluate theories behind the origins of first impressions it is important to consider what characteristics first impressions share. When reviewing the consequences of, and theories behind, first impressions some key attributes emerge. Three such attributes often reported in the literature are the: universality, automaticity, and early development of first impressions. This section will review work in each of these areas to elucidate the link between theories behind first impressions and their characteristics.

Universality

Assessing research on the universality of first impressions is one promising way through which we could tease apart evolutionary and cultural learning accounts (Nielsen & Haun, 2016; Over et al., 2020; Over & Cook, 2018). Evidence of high cross-cultural agreement in first impressions would be a strong argument that first impressions have a genetic basis routed in our evolutionary history (Sutherland et al., 2018; Zebrowitz et al., 2012). On the other hand, evidence of substantial and predictable cultural differences would support the notion that first impressions are learned culturally (Over et al., 2020).

One cross-cultural piece of work, exploring this question of universality, compared the impressions of participants from the US and the Tsimane’ people, a remote community from the Bolivian rainforest (Zebrowitz et al., 2012). When comparing the first impressions of both groups, significant within-culture agreement was found for own-culture faces on trait judgements such as sociability, dominance, and intelligence. When rating other-culture faces, significant within-culture agreement remained in both groups for judgements of intelligence (Zebrowitz et al., 2012). Similarly, research comparing US and Chinese trait judgements revealed significant cross-cultural agreement when judging own and other race faces (Albright et al., 1997). More recent research has confirmed that there are similarities between Western and Asian first impressions in a different way. Work by (Walker et al., 2011), found that participants from both cultures were sensitive in changes to other-culture faces that had been manipulated to appear more or less aggressive, extrovert, likeable, risk seeking, socially skilled, or trustworthy.

This cross-cultural work could be used as strong evidence towards an evolutionary account of first impressions, given the level of consensus seen between cultures. However, each of these studies also found evidence of cultural variability. For example, (Zebrowitz et al., 2012), when comparing Tsimane and U.S. participant’s ratings found that within-culture agreement was consistently higher in U.S. participants. One explanation for this finding offered by the authors is that U.S. participants are more likely to be exposed to media images which reinforce certain face-trait associations, a sentiment shared by TIM and cultural learning theories (Over & Cook, 2018). It was also found that Tsimane’ ratings of U.S. participants failed to reach significant within-culture agreement for measures of dominance and warmth (Zebrowitz et al., 2012). Further evidence of cross-cultural variability is shown in the work of Walker et al., (2011), who found that although Asian and Western participants were sensitive to faces manipulated on several traits, Asian participants both took longer to form impressions and, compared to Western participants, were less consistent (see also; Song & Vul., 2020).

Although Albright and colleagues (1997), found a significant relationship between US and Chinese first impressions they also found interesting differences in the cues used to form these decisions. It was found that Chinese participants strongly associated attractiveness with the trait of intelligence, whereas US participants associated attractiveness with extraversion. Additionally, amongst Chinese participants smiling was associated with a lack of self-control or calmness, a pattern not found amongst participants from the US (Albright et al., 1997). In a similar vein, when comparing Israeli and Japanese judgements of trustworthiness, researchers found that culturally specific cues, such as the typicality of a face, were key to trustworthiness judgements (Sofer et al., 2017). Differences between Japanese participants first impressions have also been compared to those of European and Asian-Americans, with results indicating that although trait inferences occur in both cultures they are both more frequent and automatic in Americans (Shimizu et al., 2017). These results, alongside the other research reviewed here, present a mixed view on the universality of trait inferences, providing evidence for a degree of universality alongside substantial cultural variability in trait judgements, not fully explained by an innate mechanism.

Another way to explore the universality of first impressions is to look at both within-culture similarities and variability in judgements. Clues to how variable first impressions can be within the same culture can be seen when we look at responses to cultural cues (Over et al., 2020). For example, some Mursi women in Ethiopia wear lip plates, a symbol of ingroup membership and positive evaluation from their peers (Turton, 2004). However, within the same community younger Mursi women are increasingly choosing not to adorn lip plates, viewing them as negative and a sign of ‘backwardness’ (LaTosky, 2006). Similar to how these cultural cues are seen differently within the same community, so may naturally varying physical cues. In fact, even those within the same household may show measurable differences in their first impressions. For example, using twins to assess the contribution of environmental and genetic factors on trust and dominance evaluations, researchers found that differences in impression formation were driven by environmental factors (Sutherland et al., 2020a). Recent research supports this assertion that environment, or top-down factors, play an important role in making trait inferences (Hehman et al., 2017). Hehman and colleagues (2017) found that for appearance-based appraisals (e.g. youthfulness, gender & race typicality) there was relatively high agreement and so responses were driven by physical cues present in the target. However, traits which require greater inference (e.g. intelligence & competence), were driven by the perceiver and so were more dependent on individual experience than other judgements.

Overall, the work on the universality of first impressions provides evidence for a degree of cross-cultural agreement. However, it is not clear whether this universality is driven by an innate, adaptive system to detect the traits of others or whether cross-cultural similarities seen in some judgements are based off salient cues to emotional expression which may convey accurate information about a target’s current, but temporary, disposition (Albright et al., 1997). Either way, wherever culturally stable judgements are found so is substantial cultural variability in trait judgements, not fully explained by an innate mechanism.

Automaticity

Consistent findings of the automaticity of first impressions, as shown by their speed and mandatory nature, represents one argument sometimes used to support evolutionary theories. The reasoning behind this is that it is the profile expected of a mechanism favoured by natural selection; for example, if someone has aggressive or nefarious intentions, it serves an organism well to detect those intentions quickly (Schaller, 2008; Zebrowitz & Zhang, 2011). The spontaneous nature of first impressions has often been cited as the ability to make trait inferences without intention or conscious awareness (Brown & Bassili, 2002; Uleman et al., 1992; Zhang & Wang, 2013). One way researchers have demonstrated this is through the speed of first impressions, showing that viewing a face for as little as 100ms is adequate time for an impression to form (Ballew & Todorov, 2007; Todorov et al., 2009; Willis & Todorov, 2006). Any extra time was shown to just increase a participants' confidence in their initial judgement. Follow-up work, with better masking procedures, have gone on to show that these judgements can occur as early as 33ms after viewing a target (Todorov et al., 2009). Evidence on the speed of first impressions demonstrate how quickly the cues, through which first impressions are formed, can be processed. Whilst speed alone implies a certain level of automaticity other work has focused on this question more directly.

The automatic nature of first impressions has been investigated using multiple different paradigms. One such paradigm is the false-recognition task, demonstrating the ease through which people bind traits to faces. Participants were more likely to falsely remember that a trait word was paired with a face if that face had previously been paired with a sentence containing a consistent trait implying behaviour (Todorov & Uleman, 2002). For example, "He threatened to hit her unless she took back what she said," may lead a participant to falsely remember that the trait term aggressive was paired with the target. To investigate the "spontaneous" nature of this phenomena manipulations were made to "deplete attentional resources". Three such manipulations used were: decreasing speed of stimuli presentation (2 seconds), increasing cognitive load (participants repeated a six number sequence), and promoting shallow processing (participants were asked to count the nouns in each behavioural description). Findings remained the same even with these manipulations (Todorov & Uleman, 2003). A similar paradigm is used within a savings in relearning paradigm (Carlston et al., 1995; Carlston & Skowronski, 1994). In this paradigm participants are again shown photos of targets and trait implying behaviours, but then after a delay are asked to learn photo-trait term pairs. Even after a two day delay it was shown that participants found it easier to learn old photo-trait pairings rather than new photo-trait pairings (Carlston & Skowronski, 1994). Findings suggest that in the initial learning phase traits were automatically bound to the target despite no trait being present in the behavioural description. This is despite no recollection of the inferred trait (Carlston et al., 1995) or memory for the initial behaviour (Carlston & Skowronski, 1994). Both the false recollection and savings in relearning paradigms demonstrate the efficiency of binding traits with faces, a process which seems to require very little cognitive resources, one key characteristic of spontaneous processes (Bargh, 1994).

Another way through which the automaticity of first impressions has been explored is through asking participants to ignore certain cues (Alguacil et al., 2015; Mileva et al., 2017; Ritchie et al., 2017; Thierry et al., 2021). This reflects their mandatory nature and how difficult certain cues are to inhibit, an aspect of first impressions explored in Chapter 2. For example, in the domain of emotional expression, participants were unable to ignore the expression of a target when asked to during an economic trust game (Alguacil et al., 2015). Even though the expression of the target did not accurately reflect their behaviour, Alguacil and colleagues (2015) still found an effect of emotional interference in participants trusting behaviour. More specific to first impressions, it has been found that when asked to ignore facial cues to dominance, and instead focus on vocal cues, participants fail to ignore facial cues fully (Mileva et al., 2017).

Researchers remain sceptical that these aspects alone constitute a strong argument for innateness. One such counter-argument draws parallels between spontaneous trait inferences and the Stroop Effect. Stating that the Stroop Effect, dependent on a learned skill (reading), also occurs quickly and automatically (Over & Cook, 2018). Furthermore, the automatic propagation of face-trait mappings is a key aspect of cultural learning models such as TIM. It is this argument which is explored in more detail in Chapter 2. Research also suggests that there may be key differences in the speed and automaticity of trait judgements across cultures (Shimizu et al., 2017; Walker et al., 2011), weakening arguments that first impressions constitute a universal preparedness. Overall, multiple streams of evidence have demonstrated the speed and efficiency of first impressions, as well as the inability to inhibit them. These all indicate the spontaneous nature of first impressions, but in isolation to other characteristics are not too revealing when considering theories regarding their underlying mechanisms.

Early Development

Previously a relatively neglected area of the literature, a surge of developmental research has increased our understanding of the processes behind trait inferences in recent years. Research had already hinted at the early emergence of children's categorisation abilities based on demographic variables, with children from at least the age of five categorising people on the basis of their race and gender (Dunham et al., 2015, 2016; Shutts et al., 2011). The work of Dunham and colleagues (2015) is particularly revealing as it demonstrates that four-year olds' race judgements were almost solely based on skin colour, whereas by age six and into adulthood, physiognomic features were increasingly relied upon.

Research on children's sensitivity to physiognomic cues in relation to trait evaluations reveals a similar pattern, and this early development has been interpreted by some as evidence that first impressions are governed by an innate mechanism (Cogsdill et al., 2014; Jessen & Grossmann, 2016). Multiple studies have found that children as young as four show high levels of consistency in attributing traits such as trustworthiness, dominance and competence to faces based only on physical facial cues, with this reaching adult like levels by 5-6 years (Antonakis & Dalgas, 2009; Cogsdill et al., 2014; Palmquist et al., 2019). In a forced choice task in which children were asked, "Which of these adults/children/monkeys is very nice?" children aged 3-5 years correctly picked the face previously rated by adults as nice at an above chance level for all face types (Cogsdill & Banaji, 2015). The strength of these physiognomic cues on children's (5 years old) and adults' trait inferences have further been shown by comparing judgements based on face-trait and face-race cues. It was found that, regardless of age, participants prioritised face-trait cues over face-race cues when making decisions on characteristics such as trustworthiness, submissiveness and competence (Charlesworth & Banaji, 2019).

In line with the adult literature, the trait impressions made by children also result in measurable behavioural consequences. Research using an economic trust game, adapted for children, showed that children aged 5 and 10 were more likely to invest resources in individuals who appear trustworthy (Ewing et al., 2015). Results mirrored in work investigating 5-year-olds pro-social behaviour in a resource management task (Charlesworth et al., 2019). Evidence from studies such as those presented above have fed the theory that children possess an adaptive mechanism to identify niceness/trustworthiness, which from an evolutionary perspective may have been used to identify those likely to do us harm and so aid in survival (Collova et al., 2020a). Furthermore, combining developmental and cross-cultural work reveals that the ability to form these trait judgements from early childhood is not culturally specific. Work with older children (8-12 years) from China aligns with results from western samples, demonstrating children's overall ability to form first impressions and high agreement between adult and children's trustworthiness judgements, as well as the cues used to make them (Ma et al., 2015b, 2016; Wang et al., 2018).

Whilst the developmental literature clearly presents evidence that children show adult-like first impressions, it is unclear to what extent this constitutes strong evidence in support of evolutionary accounts. The strength of this argument is tested directly in Chapter 2, looking at children's sensitivity to culturally learned cues, but previous literature also sheds some light on the potential role of a learning mechanism over an innately specified one. For example, one argument argues that consistency in first impressions across age groups is evidence that extended cultural learning is not a necessary prerequisite for trusting behaviour, as otherwise a developmental increase in trait biases would be expected (Ewing et al., 2019). However, in line with cultural learning accounts (Lee et al., 2021; Over & Cook, 2018), differences can be seen throughout development. For example, Mondloch et al (2019) find that for both subtle and explicit displays of emotion children do not display adult-like patterns of behaviour. In situations requiring either a trustworthy or dominant partner, children’s partner selection in contrast to adult responses, did not reveal a combined influence of facial expression and target trait.

Further supporting an increase in the use of first impressions with age, other work has suggested that there are age related differences in the way children use trait inferences to guide their behaviour (Charlesworth et al., 2019). In assessing children’s face-trait behaviour in relation to a larger number of traits than previous research, and from a younger age, Charlesworth and colleagues (2019), found that whilst consistent trait evaluations could be made by 3-year-olds, significant age differences could be seen in behavioural tendencies. Both children’s ability to use facial features to predict a target’s behaviour and to guide their own behaviour changed with age, with neither skill emerging until 5 years of age. This compliments other recent work showing that whilst four-year-olds seemed to rely on the holistic processing of face valence, five-year-olds were able to evaluate a target based on the trait apparent in each face (Palmquist & DeAngelis, 2020). This increase in adult-like first impressions with age is in line with cultural learning models suggesting that first impressions are a result of accumulated face-trait experience (Li et al., 2021) and a refinement of these mappings into adulthood (Lee et al., 2021; Over et al., 2020; Over & Cook, 2018).

Research investigating trait inferences in infancy is scarce compared to that completed with older children. Work on infants’ early experience with faces has recorded that infants are exposed to human faces for around 25% of their waking time (Sugden et al., 2014). Based on these early experiences infants are able to make surprising connections from nine months, such as associating own-race faces with happy music and other-race faces with sad music (Xiao et al., 2018). Evidence of early learned associations are evident within this previous literature (Kinzler et al., 2007; Moon et al., 1993), suggesting that sensitivity to cues in infancy can be the product of early learning opportunities over innately specified capabilities. However, there is no doubt that infants are sensitive to some aspects of face perception from birth, demonstrated by new-borns' preference for face like stimuli (Johnson, 2005; Johnson et al., 1991). Therefore, when considering theories behind first impressions, evidence of mature face-trait associations in infancy may provide strong evidence for nativist accounts (Zebrowitz & Zhang, 2011).

Work assessing infants' ability to make trait judgements has provided behavioural evidence and demonstrated different ERP responses to faces variant in trustworthiness and dominance, suggesting that by 7 months of age infants are sensitive to changes in these traits (Jessen & Grossmann, 2016, 2017, 2019) providing a case for an innate mechanism. However, currently the research is confounded by subtle yet perceivable emotional cues in the stimuli used. As a result of this it is hard to disentangle whether infants are sensitive to the perceived emotional state of the target shown or their perceived character trait. Evidence for this comes from infant research which, using the same methodology as Jessen & Grossman, found that infants were only sensitive to trustworthiness if the high and low trustworthy face were both also high in dominance (Sakuta et al., 2018). Closer inspection of the stimuli used in this study reveals that the high dominance stimuli both have clearly perceivable emotional expressions whereas the low dominance faces do not, meaning the findings could be explained in terms of emotion perception rather than trait perception.

Recent work attempting to reproduce the results by Jessen & Grossman (2016), using real over computer generated faces, was only able to find a partial replication (Baccolo et al., 2021). Work by Baccolo and colleagues (2021), was not able to find evidence that infants preferentially look towards high trustworthy over neutral or low trustworthy stimuli. Furthermore, although some similarities were found in infants' ERP responses to these three types of stimuli, significant differences were also found. Baccolo and colleagues (2021), explain that some of these differences were due to a processing advantage for mostly positively valenced faces, which may be more evident in the realistic stimuli used. Indeed, it has been hypothesised that infants' and children's early ability to track and generalise emotional dispositions may be a necessary precursor to mature trait attributions (Baccolo & Macchi Cassia, 2020; Repacholi et al., 2016). A mechanism such as one that is sensitive to salient emotional cues (e.g. smiling) is analogous to that of the innate mechanism suggested to play a small role within the cultural learning model of first impressions (Over & Cook, 2018).

Overall, work on the early development on first impressions is undoubtedly important. Early work suggests adult-like impressions are formed in early to mid-childhood (Antonakis & Dalgas, 2009; Cogsdill et al., 2014; Palmquist et al., 2019) and that these impressions interact with children's behaviour (Charlesworth et al., 2019). This may have important consequences on early childhood experiences such as the amount of pro-social directed behaviour directed towards children with trustworthy physical features compared to untrustworthy features (Charlesworth et al., 2019). This in turn may have an impact on the target's actual trustworthiness through means of a self-fulfilling prophecy (Collova et al., 2021). The extent to which these behaviours are governed by an innate versus cultural learning mechanism are not yet clear but the implications of each are important to consider. Evidence supporting a cultural learning mechanism would provide promise for early interventions that could address potential harmful behaviours early in childhood. In contrast, an innate mechanism, considering a limited role of cultural learning, may consider the success of such interventions to be extremely limited.

The present research aims

The primary aim of my doctoral work was to answer two key questions left outstanding in the literature reviewed so far on first impressions. First, how plausible are cultural learning accounts of first impressions compared to alternative accounts? Second, how might face-trait mappings be learned early in development? In my first experimental chapter (Chapter 2), I explore this first question by investigating whether first impressions that are acquired through experience display characteristics often cited as evidence for evolutionary accounts. Across three initial studies with adults, it was tested whether impressions from cultural cues (glasses) 1) can manifest consistently within adults; 2) can occur automatically, i.e., involuntarily; 3) and can be seen following brief stimulus exposure. The rationale being that if the speed and automaticity of first impressions can be observed from learned cues it is not strong evidence that first impressions have an innate origin. In a final developmental study, I tested how sensitive children aged four & six years were to both physical and cultural cues to intelligence. Evidence of sensitivity to both types of cue would indicate that the early emergence of first impressions is not sufficient evidence to support evolutionary accounts, speaking instead to the plausibility of learning-based accounts.

In my next empirical chapter (Chapter 3) I begin to explore *how* first impressions might be learned early in development as well as why these impressions might be widely shared yet often inaccurate. Previous developmental work has demonstrated children's sensitivity to non-verbal cues and how this influences their social knowledge and decision making (Brey & Shutts, 2015; Nielsen, 2013; Over & Carpenter, 2009). Based on this previous work I tested the possibility that children may also learn about the traits of others through similar non-verbal cues. To accomplish this, children aged five to seven years were asked which of a pair of targets they thought was the nicest after seeing each of these targets either associated with positive (happy reaction) or negative (fearful reaction) non-verbal behaviours from peers. Additional measures were then used to assess whether, as well as learning about the target's apparent trustworthiness, this learning would generalise to individuals similar in appearance to that target. These impressions would be formed in the absence of any diagnostic or confirmatory behaviour from the target regarding their trustworthiness. As such, these studies test whether social referencing is one route through which consistent, but also inaccurate, first impressions could emerge within a culture.

A final pair of studies, presented in Chapter 4, aimed to explore another route through which first impressions may be acquired, parent and child interactions. Previous work has demonstrated the influence of culture on children's preferences (Davis et al., 2021; Engel-Yeger & Jarus, 2008) and how important interactions with parents are to a child's understanding of the social world (Guajardo et al., 2009; Pirchio et al., 2018; Shahaeian et al., 2014). Following this work, two studies were designed to assess the plausibility that parent-child conversations are one route through which the inter-generational transmission of first impressions could occur within a culture. Across two exploratory studies parent-child dyads were presented with a picture book containing images of faces, either computer generated (Study 1) or real (Study 2), and asked to engage in conversation about each of them. Of interest was whether dyads would spontaneously use trait terms when discussing each face and how often this would occur. These studies also aimed to observe whether parents would exhibit reinforcing behaviour by endorsing inferences made by their child and how discussions about each character's traits are initiated, i.e., whether conversations are parent or child led. How parents' beliefs about making first impressions from physical appearance relate to their behaviour during the storybook task is also explored.

Chapter 2: First Impressions from Glasses

Reference:

Eggleston A, Flavell JC, Tipper SP, Cook R. Culturally learned first impressions occur rapidly and automatically and emerge early in development. Dev Sci. 2021;24: e13021. https://doi.org/10.1111/desc.13021[1]

Abstract

Previous research indicates that first impressions from faces are the products of automatic and rapid processing and emerge early in development. These features have been taken as evidence that first impressions have a phylogenetic origin. We examine whether first impressions acquired through learning can also possess these features. First, we confirm that adults rate a person as more intelligent when they are wearing glasses (Study 1). Next, we show this inference persists when participants are instructed to ignore the glasses (Study 2) and when viewing time is restricted to 100 ms (Study 3). Finally, we show that 6-year-old, but not 4-year-old, children perceive individuals wearing glasses to be more intelligent, indicating that the effect is seen relatively early in development (Study 4). These data indicate that automaticity, rapid access and early emergence are not evidence that first impressions have an innate origin. Rather, these features are equally compatible with a learning model.

---

[1] The author, Adam Eggleston, designed the experiment, collected the data, analysed the results and wrote the article under the supervision of Dr. Harriet Over.

Introduction

Adults spontaneously attribute a wide range of traits to strangers based solely on their facial features. These ‘first impressions’ include judgements about trustworthiness, honesty, competence, intelligence, aggression, and likeability (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Todorov et al., 2015). These first impressions appear to exert a powerful influence over behaviour. For example, individuals who look competent are more likely to be elected to public office (Ballew & Todorov, 2007; Todorov et al., 2015). Individuals who appear untrustworthy are less likely to be offered positions in interview settings (Olivola et al., 2014a) and are more likely to face harsh sentences in criminal justice situations (Wilson & Rule, 2015). Interestingly, while some of these first impressions may contain a ‘kernel of truth’ (Bonnefon et al., 2017), others bear little resemblance to the actual personalities of the individuals being judged (Rule et al., 2013).

Despite their pervasive influence, it is unclear where these first impressions come from. One possibility is that spontaneous trait inferences from faces are the product of an innately specified mechanism. Natural selection may have favoured such a mechanism because it enabled our ancestors to quickly detect potential cooperative partners and potential leaders (Schaller, 2008; van Vugt & Grabo, 2015; Zebrowitz & Zhang, 2009).

Alternatively, first impressions may arise chiefly through learning. According to the Trait Inference Mapping (TIM) framework (Over et al., 2020; Over & Cook, 2018), first impressions are the result of mappings between points in face-space and trait-space, acquired ontogenetically as a product of correlated face-trait experience. Once acquired, these mappings allow excitation to spread automatically from face representations to trait representations. While some face-trait mappings may be acquired through direct experience with individuals, TIM proposes a central role for cultural learning. Cultural devices such as storybooks, films, television, literature, propaganda, art and iconography, pair particular facial features (e.g., big nose, crooked teeth, pallid complexion) with particular character traits (e.g., evil disposition). By providing a common source of correlated face-trait experience, these devices allow face-trait stereotypes to spread throughout a community.

It has been argued that evidence from three sources broadly favours a nativist view of first impressions, over the learning account. Firstly, spontaneous trait inferences show high-levels of consistency within Western observers, and similar first impressions have been documented in other cultures including China (Sutherland et al., 2018; Zebrowitz et al., 2012). Secondly, traits are inferred quickly and automatically, seemingly with minimal cognitive effort. Adults form consistent first impressions from faces even when those faces are presented for as little as 100 milliseconds (Bar et al., 2006; Todorov et al., 2009; Willis & Todorov, 2006). This is the profile one might expect of a mechanism favoured by natural selection; for example, if someone has aggressive or nefarious intentions, it serves an organism to detect those intentions quickly (Schaller, 2008). Finally, proponents of the nativist view point to the fact that first impressions emerge early in development (Cogsdill et al., 2014; Ewing et al., 2019; Jessen & Grossmann, 2017). For example, developmental research has demonstrated adult-like levels of consensus in first impressions of trustworthiness/niceness, dominance/strength and competence/intelligence in Western children by at least 5 years of age (Charlesworth et al., 2019). This has led authors to the conclusion that “extended cultural learning is not necessary for adult-like appearance biases to emerge” and that findings are “more consistent with evolutionary-based accounts.” (Ewing et al., 2019).

We contend that the foregoing findings, while often cited as evidence in favour of the nativist view, are in fact equally compatible with the cultural learning account (TIM; Over & Cook, 2018; Over et al., 2020). For example, exposure to correlated face-trait experience through cultural devices (e.g., story books, films, TV) may cause face-trait inferences to manifest consistently both within and between cultures. Similarly, it is known that learned behaviours can become rapid and automatic (e.g., Stroop, 1935), and can appear even in infancy (Kinzler et al., 2007; Moon et al., 1993).

In the present paper, we describe a series of experiments that sought to test directly whether first impressions that are acquired through experience: 1) can manifest consistently within adults; 2) can occur automatically, i.e., involuntarily; 3) can be seen following brief stimulus exposure; and 4) can be observed early in development. To test these hypotheses, we took advantage of the fact that participants' first impressions of a stranger's intelligence are sometimes influenced by whether or not the stranger is wearing glasses. The earliest known glasses date from 1286 (Ilardi, 2007). Because glasses are a product of recent human history, it is indisputable that, where observed, the inference of intelligence from glasses must be acquired through experience (Over & Cook, 2018). The data for all studies are available at the OSF: (https://osf.io/hkpsa/?view_only=575e24eee0e64242aa2a6525683d1ac0).

Study 1

In Study 1, we sought to replicate previous findings from cultures such as the US (Fleischmann et al., 2019) and Scandinavia (Hellström & Tekle, 1994) that individuals wearing glasses are judged more intelligent.

Method

Participants

Forty participants were recruited via the online platform Prolific (www.prolific.ac.uk). Sample size was decided in advance, and pre-registered, based on previous research in this area. All participants reported English as their first language. One participant was excluded because they selected the same rating on 100% of trials (per the exclusion criteria pre-registered via ASPredicted) leaving 39 for analyses (M_{age} = 32.51, SD_{age} = 13.59; 16 female, 22 male, 1 non-binary). Of the 39 participants included in the analyses, 37 identified as White and 2 as Multiracial. All participants received a monetary compensation of £1.30 for approximately 10 minutes participation.

Materials and procedure

Twenty-four images of White faces (12 Female) were taken from the Chicago Face Database (Ma et al., 2015a)**.** This particular database was selected as it is able to provide full colour face images of neutral expression under standardised photographic conditions, including consistent lighting, clothing and position. Glasses were added to these stimuli using Adobe Photoshop. This resulted in 48 (24 altered, 24 unaltered) images that were combined to create two counterbalancing conditions (Figure 2). Participants were assigned to one of these two conditions so that the faces wearing glass in condition A were not wearing glasses in condition B and *vice versa*. The order in which the faces were presented was randomised for each participant.

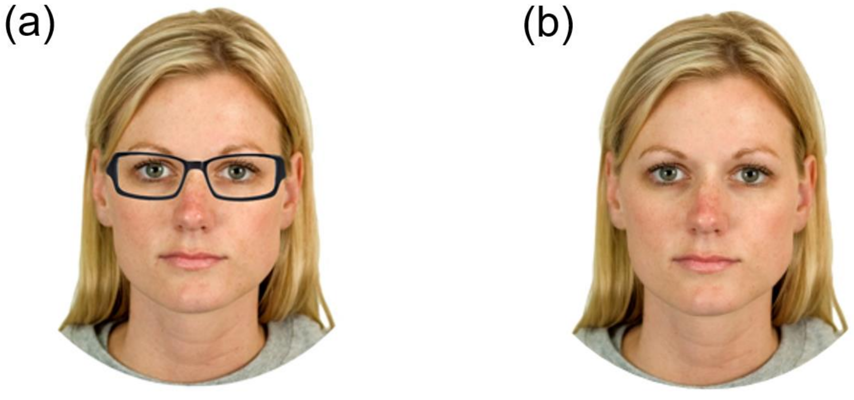

Figure 2. Example stimulus from Studies 1-3, shown with (a) and without (b) glasses. Unedited stimuli taken from the Chicago Face Database (Ma et al., 2015a)

Ratings of intelligence were gathered using an online survey designed in Qualtrics (http://qualtrics.com). Participants were informed that they would complete a study investigating how people form first impressions of others. They were the told that their task was to view photographs of people and rate how intelligent they appeared. On each trial, participants were shown a face and asked “How intelligent do you think this person is?” The task was self-paced: participants gave their responses on a 9-point Likert Scale with each face remaining on the screen until participants clicked a continuation arrow to proceed to the next trial. At the end of the procedure, participants were asked to complete a short demographic questionnaire recording gender, ethnicity, age and whether or not they had a prescription for glasses. After participants had completed all parts of the questionnaire they were thanked and debriefed.

Results

Following our pre-registered analysis plan (https://aspredicted.org/blind.php?x=8w4fr8), ratings of intelligence (Figure 3a) were analysed using a Mixed ANOVA with Face Type (glass, no glasses) as a within-subjects factor, and Counterbalancing Condition (set A, set B) as a between-subjects factor. The analysis revealed a significant main effect of face type [$F(1,37) = 93.08$, $p < .001$, $\eta p^2 = .72$], whereby glasses wearers ($M = 6.31$, $SD = 0.74$) were rated as more intelligent than those without glasses ($M = 5.12$, $SD = 0.82$). There was no main effect of Counterbalancing Condition [$F(1,37) = .05$, $p = .819$, $\eta p^2 = .001$], nor did we observe an interaction between Counterbalancing Conditions and Face Type [$F(1,37) = .07$, $p = .793$, $\eta p^2 = .002$]. All *p*-values described throughout are two-tailed.

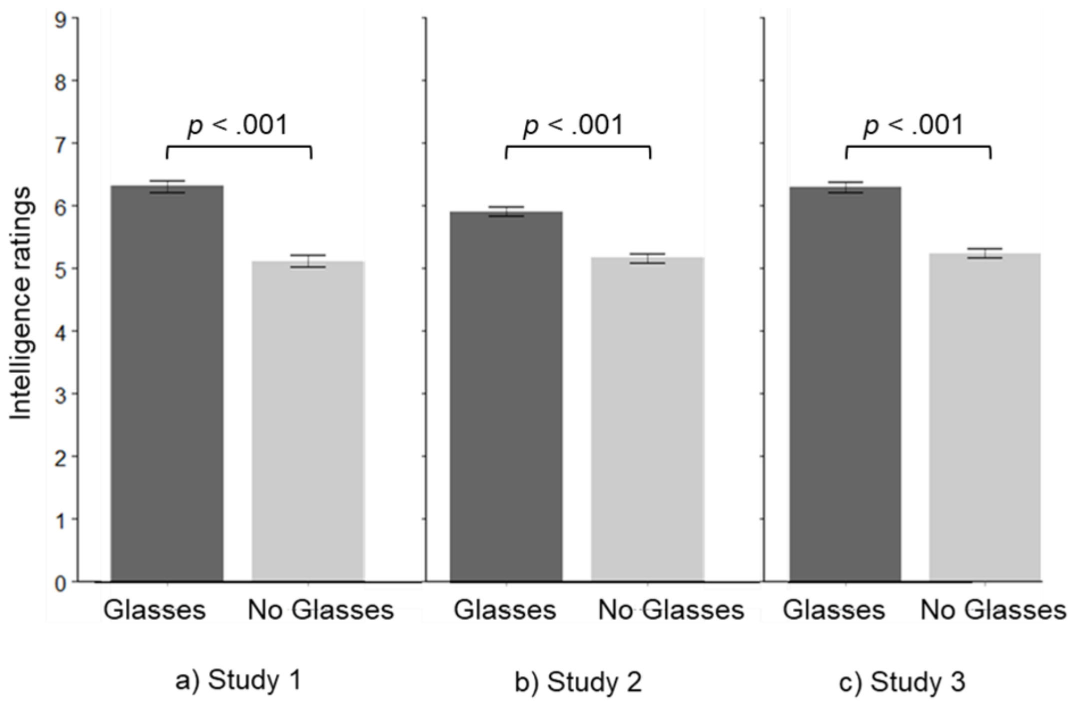

Figure 3*.* Mean (±SEM) intelligence ratings Studies 1-3.

Study 2

In Study 2, we tested whether first impressions of intelligence are inferred automatically from the presence of glasses. If the inference of intelligence is automatic, the glasses effect should still be seen when participants are explicitly instructed to ignore the glasses and base their judgements only on physical facial cues (Alguacil et al., 2015; de Gelder & Vroomen, 2000; Mileva et al., 2017).

Method

Participants

A further forty participants were recruited via www.prolific.co.uk following our pre-registered plan (M_{age} = 34.93, SD_{age} = 9.97; 27 female, 12 male, 1 non-binary). Thirty-seven identified as White, 1 as Asian, 1 as Black, and 1 as "Other". No participants were excluded from analysis.

Materials and procedure

Participants were given the following instructions before starting the trials; "You are going to see a series of photographs of different people. Some of the individuals depicted will be wearing glasses and some will not. Glasses can make people look more intelligent. We want you to focus only on the structure of the face.
Please ignore the glasses." Where trials presented a target face wearing glasses, participants also received the following reminder: "Remember: Ignore the glasses." As a check to see if participants had been paying attention throughout a final question was added asking "What feature were you asked to ignore?" Participants could respond by typing their answer (all but one participant correctly identified

glasses as the cue they were supposed to ignore). With the exception of the instruction to ignore the glasses, the procedure was identical to that of Study 1.

Results

Following our pre-registered analysis plan (https://aspredicted.org/blind.php?x=y5i547), ratings of intelligence (Figure 3b) were analysed using ANOVA with Face Type (glass, no glasses) as a within-subjects factor, and Counterbalancing Condition (set A, set B) as a between-subjects factor. The analysis revealed a significant main effect of Face Type [$F(1,38) = 61.73$, $p < .001$, $\eta p^2 = .62$], whereby glasses wearers ($M = 5.91$, $SD = 0.57$) were rated as more intelligent than those without glasses ($M = 5.17$, $SD = 0.69$). There was no main effect of Counterbalancing Condition [$F(1,38) = .02$, $p = .892$, $\eta p^2 = < .001$], nor did we observe an interaction between Counterbalancing Conditions and Face Type [$F(1,38) = .95$, $p = .336$, $\eta p^2 = .02$].

Study 3

In our first two studies, we replicated the glasses effect (Study 1), and found that the effect is seen despite explicit instructions to ignore the glasses (Study 2). In both studies, the task completed by participants was self-paced; i.e., participants were free to inspect target faces for as long as they wished. In Study 3, we sought to determine whether the glasses effect would be seen when stimulus presentation was limited to 100ms. Whereas the first two studies were conducted online, Study 3 was conducted in the lab.

Method

Participants

Forty participants were recruited from the University of York Psychology Department subject pool (M_{age} = 20.39, SD_{age} = 4.72; 34 female, 6 male). Sample size was decided in advance, based on the results of the first two studies. One participant declined to specify their age. Thirty-two participants identified as White, 2 as Multiracial, 1 as Indian, 4 as British (not otherwise specified) and 1 preferred not to say. Participants received course credit or a small honorarium for taking part.

Materials and procedure

At the start of each trial, a fixation cross appeared in the centre of the screen for 500 ms followed by 500 ms of blank screen. A target face then appeared for 100 ms before being masked for 500ms (Figure 4a). The rating screen appeared after 1000 ms of blank screen. Participants then rated the face for intelligence on a 9-point Likert scale (Figure 4b) using the mouse. The stimuli used in Study 3 were the same as those used in Studies 1 and 2. Faces were presented on a white background with 28mm distance between pupils. The square mask had sides of ~147mm.

Before completing the main experiment, participants first completed 4 practice trials. Participants then completed 24 experimental trials in a single block. Before the practice block participants were given on-screen instructions stating that they would see people presented for a very short period of time and that they would then rate how intelligent they thought the presented people were. Additionally, they were told that they should make their decisions as quickly as possible. At the end of the experiment participants completed a short questionnaire (age, gender, ethnicity, first language and glasses prescription).

The task was completed in a dimly-lit room on a Dell (Round Rock, USA) XPS PC (Intel Core 133 i5-4430, 3 GHz CPU, 12 GB RAM, 64 bit Windows 10 Enterprise), with a 23" LCD monitor (Iiyama (Tokyo, Japan) ProLite T2735MSC-B2, 1920×1080 pixels). Viewing distance was approximately 50 cm distance. Stimulus presentation (60Hz) and response recording were achieved using custom scripts and Psychtoolbox 3.0.11 (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) operating within Matlab R2018a (The MathWorks Inc., Natick, USA).

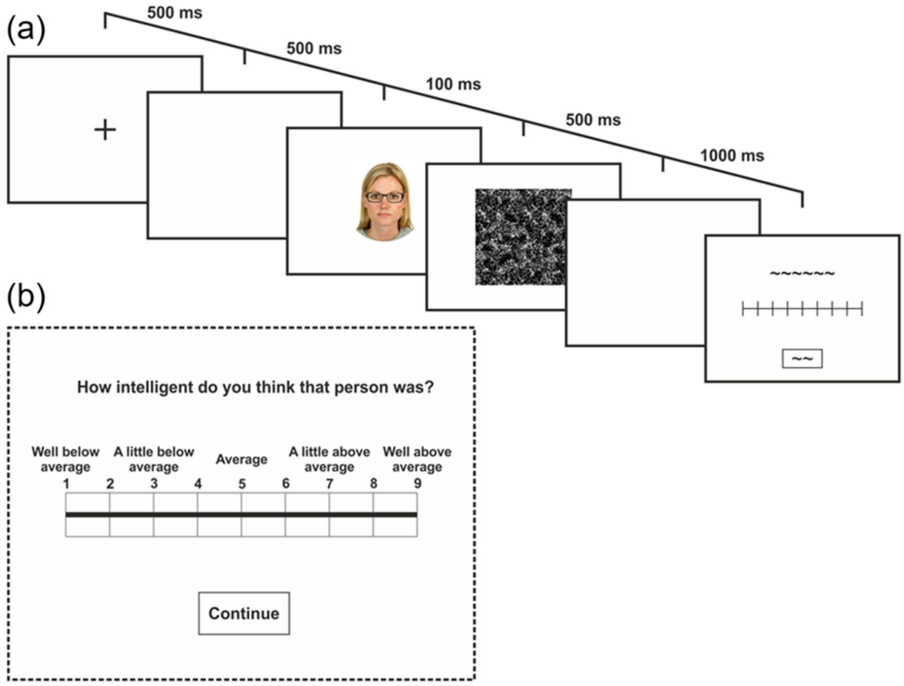

Figure 4. Trial structure in Study 3. (a) Schematic representation of a trial, (b) the rating screen.

Results

Following our pre-registered analysis plan (https://aspredicted.org/blind.php?x=r8kh3a), ratings of intelligence (Figure 3c) were analysed using ANOVA with Face Type (glass, no glasses) as a within-subjects factor, and Counterbalancing Condition (set A, set B) as a between-subjects factor. The analysis revealed a significant main effect of Face Type [$F(1,38) = 88.81$, $p < .001$, $\eta p^2 = .70$], whereby glasses wearers ($M = 6.30$, $SD = 0.64$) were rated as more intelligent than those without glasses ($M = 5.25$, $SD = 0.73$). There was no main effect of Counterbalancing Condition [$F(1,38) = .69$, $p = .412$, $\eta p^2 = .02$], nor did we observe an interaction between Counterbalancing Conditions and Face Type [$F(1,38) = .18$, $p = .670$, $\eta p^2 = .01$].

Study 4

Previous research has demonstrated adult-like levels of consensus in first impressions of trustworthiness (niceness), dominance (strength), and competence (intelligence) in Western children of ~5 years of age (Charlesworth et al., 2019; Cogsdill et al., 2014; Ewing et al., 2019) In Study 4, we sought to determine whether inferences of intelligence from glasses emerges at a similar point in development. We chose to test 4- and 6-year-olds because these age groups have been studied in previous research (Cogsdill et al., 2014) and research suggests that children's understanding of specific traits such as a person's intelligence or knowledge develops significantly across this age range (Brosseau-Liard & Birch, 2010; Lane et al., 2013).

Method

Participants

Thirty-two 4-year-olds (16 boys, M_{age} = 53 months; age range = 48 to 59 months) and 32 6-year-olds (16 boys, M_{age}: 78 months; age range = 73 to 83 months) participated in this study. Sample size was decided in advance and is typical for research in this area (Caulfield et al., 2016; Ewing et al., 2015; Mondloch et al., 2019; Palmquist et al., 2019). All 64 families who participated identified as White British in ethnicity. All children were recruited from a science museum in an urban centre and were tested on site the same day.

Materials and procedure

The stimuli in this study were again taken from the Chicago Face Database (Ma et al., 2015). Having selected 50 faces from the database (25 white female, 25 white male), we ran a pre-test in which we asked 20 undergraduate participants from the Department of Psychology at the University of York to rate the perceived intelligence of each. Ratings were taken on a scale that ranged between “0 (Not at all clever)” to “100 (Very clever)”. Participants received course credit in return for their participation.

The pre-test ratings were used to create 24 pairs of faces. The ratings for each pair can be found on the OSF. Twelve pairs of faces (6 male pairs, 6 female pairs) were chosen because adults judged them be closely matched in terms of their perceived intelligence (Figure 5a). These pairs were used to assess whether children infer intelligence from glasses. Differences in adults’ ratings of the perceived intelligence of the chosen face pairs with glasses ranged from .10 and 3.15 (M=1.55, SD = 1.01). Glasses were superimposed on the faces using Adobe Photoshop. We counterbalanced which face was shown wearing glasses. The remaining 12 pairs (6 male pairs, 6 female pairs) were chosen because adults consistently judged one of them be more intelligence than the other (Figure 5b). These pairs were used to assess whether children infer intelligence from the physical cues present in faces alone. Difference in adults' ratings of intelligence between individuals within each pair of faces used on the physical trials ranged between 4.55 and 26.30 (M=13.67, SD = 5.60). A book was created featuring the 24 to-be-judged pairs. The pairs were presented in a fixed order chosen at random before testing. High-intelligent and glasses wearing faces represented the target faces and were presented equally often on the left and right side.

Each participant was invited into the testing area and asked to sit at a small table. After a brief warm up, the experimenter tested the child's understanding of the term 'clever' by saying: "Now I am going to tell you about two different children and I would like you to tell me which child you think is the more clever, okay? This child got all their answers right on their homework and the teacher said it was really good work. This child got only a few answers right on their homework and the teacher said it wasn't good work at all. Who do you think is more clever, this child who did all their homework or this child who did half the homework?" All 64 children correctly identified the child who "got all their answers right on their homework" as "more clever". Children in both age groups therefore appear to understand the concept of relative intelligence as operationalized in the present task.

Children were then introduced to the first of the 24 face pairings by the experimenter as follows: "Now we are going to look at some pictures of different people. I would like you to tell me which person you think is more clever, okay? Take a look at these two people, which person do you think is more clever - this person or this person?" [Pointing to each picture in turn, left to right]. All remaining trials were identical except that the experimenter only pointed to the two faces in turn as a prompt if children showed hesitancy.

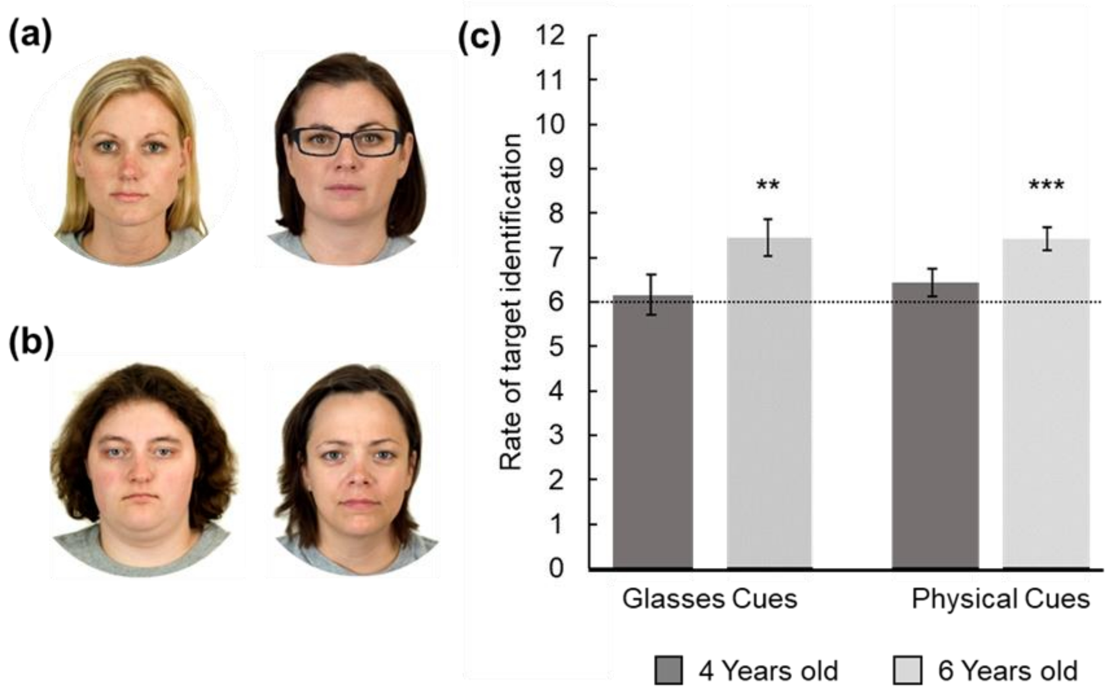

Figure 5. Stimuli and results from Study 4. (a) Example pair from the glasses task. (b) Example pair from the physical cues task. (c) Mean (±SEM) rate of correctly identifying the target face for the 4 and 6 year old children in the glasses and physical cues tasks. Dashed line indicates chance performance. Asterisks represent a result significantly different from chance. ** denotes $p < .01$; *** denotes $p < .001$.

Coding

Children's responses were coded from the video. On the rare occasions when a child changed their response, their last response was coded. For each trial type (physical cues, glasses cues), participants were given a score out of 12 for the number of times they chose the individual in glasses or the individual previously rated by adults as more intelligent.

All videos were coded separately by the experimenter and a rater naïve to the hypotheses of the study. There was perfect agreement between the two coder's judgements, $\kappa = 1$, $p < .001$.

Results

First, we sought to determine whether the children's identification of the target face exceeded chance levels when inferring intelligence from glasses and physical cues (Figure 5c). To this end, we conducted one-sample *t*-tests evaluating the number of times children correctly chose the target face, against a chance level of 50% (i.e., a score of 6 out of 12). In the glasses condition, the rate of target identification exhibited by the six-year-olds (M = 7.41, SD = 2.33) was significantly greater than chance [$t(31)$ = 3.42, p = .002, Cohen's d = .61]. However, the rate of target identification exhibited by the four-year-olds (M = 6.13, SD = 2.50) did not exceed chance [$t(31)$ = .283, p = .779, Cohen's d = .05]. A very similar pattern emerged with children's judgements of physical cues to intelligence. Once again, the rate of target identification exhibited by the six-year-olds (M = 7.38, SD = 1.48) exceeded chance [$t(31)$ = 5.27, p = <.001, Cohen's d = .93], whereas the rate of target of identification exhibited by the four-year-olds did not (M = 6.41, SD = 1.76) [$t(31)$ = 1.31, p = .201, Cohen's d = .23].

Next, we analysed the number of times the target faces were chosen using ANOVA with Cue Type (glasses cues, physical cues) as a within-subjects factor, and Age (4-year-olds, 6-year-olds) as a between-subjects factor. The analysis revealed a main effect of Age, with 6-year-olds choosing the target faces more often than did 4-year-olds [$F(1,62)$ = 8.73, p = .004, partial η^2 =.12]. However, there was no main effect of Cue Type [$F(1,62)$ = 0.13, p = .719, partial η^2 =.002] and no interaction between Age and Cue Type [$F(1,62)$ = 0.21, p = .653, partial η^2 =.003]. These results indicate that older children were more likely to identify the target faces than younger children, but that the developmental pattern was broadly similar for physical and cultural cues to intelligence.

The finding that four-year-olds do not show significant sensitivity to either cultural or physical cues to intelligence seems inconsistent with previous findings which report adult-like consistency in trait judgements of competence by this age (Charlesworth et al., 2019). This may be because previous developmental work has used computer generated stimuli whereas we used photographs of real people perhaps making the task more challenging.

General Discussion

We began our investigation by confirming that adults attribute intelligence to individuals who wear glasses (Study 1). We went on to extend the basic glasses effect by demonstrating that the inference of intelligence is seen despite explicit instructions to ignore the glasses cue, suggestive of automaticity (Study 2), and is evident following even very brief (100 ms) stimulus presentation (Study 3). Finally, we tested young children's propensity to infer intelligence from the presence of glasses and physical facial cues and found that, whereas six-year-olds use both glasses and physical cues, four-year-olds do not use either type of cue (Study 4).

Consistent impressions of trustworthiness and dominance appear to be automatic and rapid (Bar et al., 2006; Todorov et al., 2009; Willis & Todorov, 2006), and emerge early in development (Charlesworth et al., 2019; Cogsdill et al., 2014; Ewing et al., 2019). These findings have been cited as evidence that spontaneous trait inferences have an innate origin (Zebrowitz & Zhang, 2009). However, the present results demonstrate that the logic of this argument is flawed. Critically, the inference of intelligence from glasses – an effect that must emerge from learning – is automatic and rapid, and also emerges relatively early in development. Our results demonstrate that findings previously thought to support nativist accounts of the origins of first impressions are equally compatible with a learning-based account (Over & Cook, 2018).

The results described here accord with findings in other areas of cognition showing that learned skills, such as reading can become fast and automatic with sufficient experience (Heyes, 2018; Stroop, 1935). They are also compatible with research from developmental psychology demonstrating that learned preferences can appear early in development. For example, even infants prefer speakers of their native language to speakers of a foreign language (Kinzler et al., 2007).

What evidence remains for nativist accounts of first impressions? One possible source of evidence is data suggesting that even 6-to-8-month-old infants prefer to look at faces previously rated by adults as trustworthy (Jessen & Grossmann, 2016; Sakuta et al., 2018). Convincing evidence that infants under the age of one-year-old form spontaneous first impressions would support a poverty of the stimulus argument for at least first impressions of trustworthiness. However, these infant studies systematically confound the to-be-inferred trait with facial emotion; the faces described as “trustworthy” appear to be smiling whereas the faces deemed “untrustworthy” do not (Over & Cook, 2018). In the absence of evidence to the contrary, it is more likely that the orienting behaviours observed in infants in these studies are a response to the perceived emotional *states* of the actors shown, and have little to do with their perceived character *traits*.

Other evidence supports the claim that first impressions are culturally learned. For example, recent work has demonstrated the ability for individuals to update implicit and explicit judgements of trustworthiness by learning new behavioural information about a target (Shen et al., 2020). Cross-cultural work has further supported a large role of learning, demonstrating that first impressions of trustworthiness, dominance, and other character traits are highly culturally variable (Over et al., 2020). Furthermore, there is considerable variation within cultures depending on individuals' learning experiences. For example, on encountering faces that vary in ethnicity, some individuals form strong first impressions of differing trustworthiness, while others do not (Devine, 1989). This extensive variability is difficult to reconcile with nativist accounts (Over et al., 2020).

Across the four studies presented here, a range of methods regarding participant recruitment were used. The use of online recruitment has become increasingly popular over recent years (Sassenberg & Ditrich, 2019). One thing to acknowledge however is that this type of recruitment allows participants to pick and choose the studies they wish to take part in. Engagement with research on first impressions raises the possibility that those more likely to endorse first impressions are also more likely to take part in research behind them. Whilst those who do not endorse first impressions may instead decline to engage with research. Research does suggest that the belief that facial features are indicative of a person's character are wide-spread (Jaeger et al., 2020b). However, the possibility of differing engagement in first impression research based on these beliefs could be addressed by work collecting qualitative data regarding participant's attitudes towards first impressions alongside quantitative data.

Overall, findings that may initially appear to favour a nativist account of first impressions are, in reality, equally compatible with a learning account. On balance, we argue that the available evidence favours the view that first impressions from faces are acquired through experience. If correct, this conclusion has important applied implications. If first impressions are shaped by cultural learning, then changes to cultural products can alter the types of first impressions individuals form.

In other words, it may be possible to mitigate widely held but deleterious societal beliefs through modifying the available cultural input.

Chapter 3: First Impressions Through Social Referencing

Reference:

Eggleston, A., Geangu, E., Tipper, S. P., Cook, R., & Over, H. (2021). Young children learn first impressions of faces through social referencing. *Scientific Reports*, *11*(1), 1-8.[2]

Abstract

Previous research has demonstrated that the tendency to form first impressions from facial appearance emerges early in development. We examined whether social referencing is one route through which these consistent first impressions are acquired. In Study 1, we show that 5- to 7-year-old children are more likely to choose a target face previously associated with positive non-verbal signals as more trustworthy than a face previously associated with negative non-verbal signals. In Study 2, we show that children generalise this learning to novel faces who resemble those who have previously been the recipients of positive non-verbal behaviour. Taken together, these data show one means through which individuals within a community could acquire consistent, and potentially inaccurate, first impressions of others faces. In doing so, they highlight a route through which cultural transmission of first impressions can occur.

---

[2] The author, Adam Eggleston, designed the experiment, collected the data, analysed the results and wrote the article under the supervision of Dr. Harriet Over.

Introduction

Humans spontaneously attribute a wide range of traits to strangers based on their facial features (Oosterhof & Todorov, 2008). These first impressions include judgements about honesty, competence, intelligence, dominance, and likeability (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Todorov et al., 2015; Zebrowitz & Montepare, 2008). First impressions are thought to load on to two principle dimensions - trustworthiness and dominance (Oosterhof & Todorov, 2008). Developmental research has demonstrated that the tendency to form impressions from faces emerges early in development. By the age of 3, children make explicit judgements about how ‘nice’ and ‘strong’ a person is after viewing images of their face (Charlesworth et al., 2019; Cogsdill et al., 2014). These first impressions exert a measurable influence over behaviour. Ewing and colleagues have shown that children are more generous towards individuals who appear trustworthy (Ewing et al., 2019). Interestingly, research with adults has shown that while some first impressions contain a kernel of truth (Bonnefon et al., 2015), others bare little or no resemblance to the actual character traits of the individuals being judged (Dilger et al., 2017).

A key priority for developmental research is to understand the mechanism, or mechanisms, by which first impressions of faces are acquired (Over et al., 2020; Over & Cook, 2018). One view is that at least some first impressions are the product of an innately specified mechanism to distinguish between potential social partners (Schaller, 2008; Zebrowitz, 2004; Zebrowitz & Zhang, 2011). According to this view, determining who among many social partners was a potential collaborator or leader was so crucial to our evolutionary ancestors that natural selection favoured individuals who were predisposed to make these judgments (Schaller, 2008;

Zebrowitz & Zhang, 2011). In contrast, Trait Inference Mapping (TIM) offers a learning account in which all mappings between 'face space' and 'trait space' are acquired as a result of experience (Over & Cook, 2018). Consistent with this, research has shown that adults quickly learn who is trustworthy from receiving small amounts of behavioural information about them (Verosky & Todorov, 2010) and generalise this information to novel individuals who are somewhat similar in appearance (Verosky & Todorov, 2013).

Proponents of both theoretical positions agree that at least some first impressions must be learned (Cook & Over, 2020; FeldmanHall et al., 2018; Hackel et al., 2015; Over & Cook, 2018; Ramsey & Ward, 2020; Sutherland et al., 2020ab). Evidence in favour of this claim comes from data showing that participants form first impressions from cultural cues that are the products of recent human history. For example, children and adults from Western cultures typically judge individuals who wear glasses to be more intelligent than individuals who do not wear glasses (Eggleston et al., 2021). Other research has shown that there are systematic cultural differences in first impressions (Over et al., 2020) and that it is possible to modify pre-existing first impressions of faces with training (FeldmanHall et al., 2018; Verosky & Todorov, 2013).

To date, relatively little research has directly investigated *how* face-trait learning takes place. One particular challenge is to explain how inaccurate first impressions can emerge through learning (Todorov, 2017). If individuals learned face-trait mappings through direct experience, then consistent first impressions would not emerge at a group level because people's facial appearance rarely predicts their actual character traits (Dilger et al., 2017). TIM explains the prevalence of shared but inaccurate impressions through appealing to cultural learning (Over et al., 2020; Over & Cook, 2018). One route through which this could happen is exposure to cultural products such as storybooks, films, advertising and propaganda that pair particular facial features with character traits (Over & Cook, 2018). For example, children’s animations pair the presence of physical beauty in protagonists with positive behaviours (Klein & Shiffman, 2006; Ryan et al., 2018). Through exposure to such systematic messages, many children in the same community may acquire similar face-trait mappings.

The authors of TIM also predicted that children within a community might learn common (but inaccurate) first impressions from their caregivers (Over & Cook, 2018). Sometimes, teaching could take the form of explicit instruction. For example, verbal warnings that children should avoid individuals with a particular physical appearance. Other forms of teaching might be more implicit. Specifically, children might learn first impressions partly through social referencing (Over & Cook, 2018). That is, by attending to the non-verbal responses of others (Walle et al., 2017).

We know from previous research that adults’ non-verbal behaviour differs in systematic ways when interacting with individuals who differ in appearance. For example, Weisbuch and colleagues (2009) have shown that the non-verbal behaviour of White Americans in popular TV shows is less positive when interacting with African Americans than when interacting with other White Americans (see also: Castelli et al (2012)). Furthermore, both children and adults can acquire intergroup biases from observing the non-verbal behaviour of others (Skinner et al., 2019; Weisbuch et al., 2009). Indeed, research on social referencing suggests that even infants use the nonverbal reactions of others to decide who to approach and avoid (Repacholi et al., 2016).

Here, we extend this research to the domain of first impressions of faces and investigate whether children use the non-verbal behaviour of others, specifically similar aged peers, to infer the trustworthiness of faces. We chose to investigate the influence of peers on first impressions as previous research has suggested that peers exert a consistent influence over children's choices and preferences in social settings (Shutts et al., 2010). We test this question with 5- to 7-year-olds because we know that children in this age range form consistent first impressions on the basis of others' appearance (Charlesworth et al., 2019; Cogsdill & Banaji, 2015; Ewing et al., 2019) and engage in extensive social learning (Over & Carpenter, 2015)

Study 1

The main aim of this study was to determine whether children use the non-verbal behaviour of others to make attributions of facial trustworthiness. We presented children with computerised displays in which target faces were paired with context faces that appeared either happy or afraid. We predicted that when children were asked "who do you think is nicer" that they would choose target faces associated with happy context faces significantly more than target faces paired with fearful faces. Following previous developmental research in this area, we used the term "nice" rather than "trustworthy" as the younger children in our sample may not understand the term "trustworthy" (Charlesworth et al., 2019; Cogsdill & Banaji, 2015). We also sought to test whether children generalise from these learned associations to similar looking but novel individuals. We predicted that children would show a preference for the composites constructed from targets previously associated with happy context faces.

Method

Pre-registration and open science

Both studies were pre-registered. The pre-registered details for Study 1 (https://aspredicted.org/pn5vq.pdf) and Study 2 (https://aspredicted.org/xv6qs.pdf) are available. The data is also available open access at Open Science Framework: https://osf.io/eu4q8/?view_only=a82c1d51e83a4a2fa26b553746830b69

Participants

The final sample consisted of 120 children, with equal numbers of 5-year-olds (20 boys M_{age} = 66 months, age range = 60 to 71 months), 6-year-olds (20 boys M_{age} = 76 months, age range = 72 to 82 months), and 7-year-olds (20 boys M_{age} = 90 months, age range = 84 to 95 months). An additional four children were tested but excluded from analysis in line with our pre-registered exclusion criteria (they required more than 4 prompts to look at the screen). Of the 120 participants included in the analysis, 106 were identified by their parents as White British, 1 as White European, 1 as White Irish, 2 as British/Indian, 5 as British/Pakistani, 1 as British/Bangladeshi, 1 as White/Black Caribbean, 1 as Asian Mixed, 1 as White and Black African and 1 as Mixed English/Arab. All children were recruited from a science museum in an urban centre and were tested on site the same day. Informed written consent was gathered from a parent of every child tested and assent was gained from each child. The procedure was approved by the University of York Department of Psychology's Ethics Committee and all methods were performed in accordance with the committee's guidelines.

Materials

The stimuli were photographs of children's faces taken from the Dartmouth Database of Children's Faces (Dalrymple et al., 2013) and all faces used in the figures throughout are those for which Dalrymple and colleagues obtained assent from the child and written informed consent from the parent to both have their photographs distributed to other researchers and for the photographs to be used in scientific publications. This database was chosen as it provided high quality and constrained images of children's faces featuring a range of expressions at different angles. The ages of the children depicted in the photographs ranged from 8 – 10 years, were all white, and contained an equal number of males and females. The background was removed from each image and faces were presented on a black background. Target faces were presented facing the camera and context faces were presented in profile and appeared to observe the target face (see *Figure 6A*). In each of the 4 context pairs, one face was female and one was male.

It is possible that baseline levels of trustworthiness will affect learning and so we chose pairs of target faces that resembled each other as closely as possible in apparent trustworthiness. In order to achieve this, we ran a pre-test in which we asked 20 undergraduate participants to rate the children's faces on a 100 point slider-scale ranging from 'not at all nice/trustworthy' to 'extremely nice/trustworthy'. Target pairs were then matched based on gender and average nice/trustworthiness ratings. The relative trustworthiness of the 16 targets are available in Table A Appendix A.

Stimuli for the two generalisation trials were created using WebMorph (DeBruine, 2017) an online programme specifically created for image manipulation and transformation (see Tiddeman et al (2001, 2005) for a detailed description of the process). The generalisation stimuli are the result of averaging the 4 target faces previously paired with happy context faces and the 4 target faces previously paired with fearful context faces to create 4 composite images in total – two for the male target faces and two for the female target faces.

Design

Participants completed 8 standard trials, each with a unique pair of faces, followed by a final generalisation test (a male test and female test). Within each standard trial, there were 8 learning events; 4 in which one of the target faces acquired positive valence through context faces moving from a neutral to happy expression, and 4 in which the other target face in the pair acquired negative valence through context faces moving from a neutral to fearful expression (Figure 6A). The 8 learning events lasted 49 sec in total. At the end of each trial, participants were shown the two target faces side-by-side (see *Figure 6B*) and were asked a forced choice question: ‘Who do you think is nicer, this person or this person?’.

Half of trials included male target faces and half included female target faces. Half of trials begin with happy context faces and half begin with fearful context faces. In addition, the ‘trustworthy’ face appeared on the left in 50% of trials. The target faces associated with happy and fearful context faces were switched in two counterbalancing conditions, and within these counterbalancing conditions trials were presented in one of two possible orders. Overall, this resulted in four between subjects counterbalancing conditions to which participants were randomly assigned.

Procedure

Participants were invited into the testing area and asked to sit at a small table in front of a laptop. After a brief warm up, the experimenter (E) conducted a comprehension test to assess children's understanding of the term nice. To do this E presented a pair of male faces and said: "First I am going to tell you about two different people. This person shared a cookie with another person in their class and this person stole a cookie from another person in their class. Who do you think is nicer? This person or this person?" [Pointing to each picture in turn]. Children were corrected if they chose incorrectly and a note made of their decision.

The first of the experimental trials began with E saying: "Now I'm going to show you some more people, a bit like these, and it's going to play a bit like a video. At the very end of the video I am going to ask you again who you think is the nicest, okay? Please try and look at the screen for the whole video." E then started the presentation of the first trial. Following the onset of the two target faces (side-by-side), E asked, 'Who do you think is nicer – this person or this person?' [Pointing to each picture in turn]. The same procedure and wording was used for all remaining trials. To ensure children stay engaged with the task, after every two trials there was a short break where children were offered a sticker to add to their bookmark. The two generalisation trials followed immediately after the final test trial with no break in between. Children were presented with two more pairs of target faces (one male pair and one female pair) and were asked again 'Who do you think is nicer – this person or this person?' [Pointing to each picture in turn].

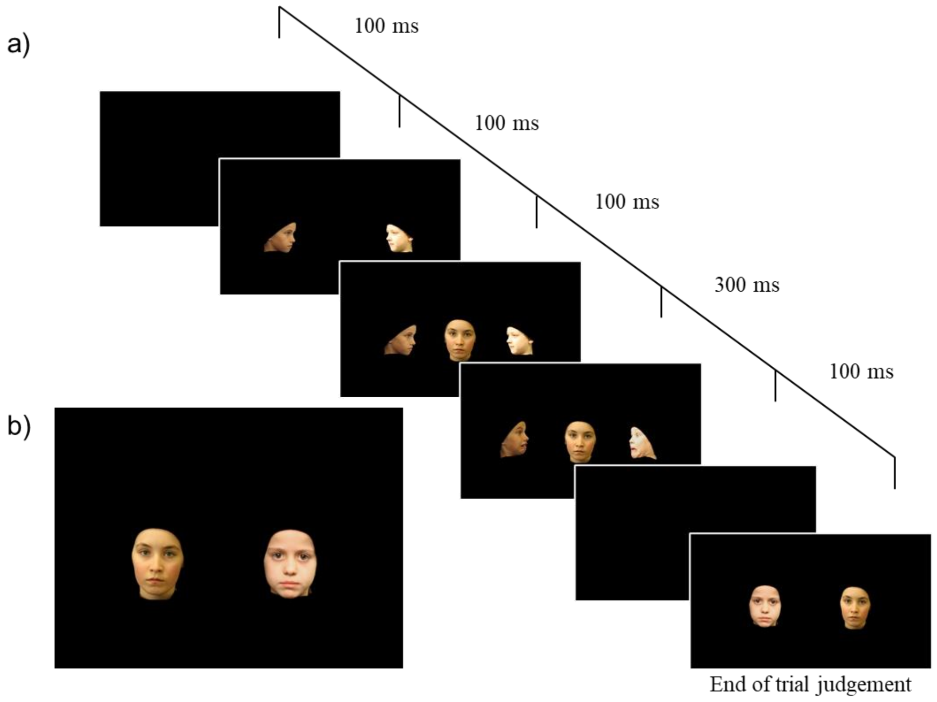

Figure 6. Stimuli in Study 1. (A) Schematic representation of frames from a learning event incorporating fearful context faces (B) Example test trial.

Coding

Participants were given a score out of 8 for the number of times they chose the target faces that had been paired with happy context faces. For the generalisation trials, children were given one score for the female target pair and one score for the male target pair. 25% of the data were second coded by a rater who did not know the hypotheses of the study. There was perfect agreement between the two coder's judgements, κ = 1.

Results

Following our pre-registered analysis plan, we conducted a one-sample *t*-test evaluating the number of times the children (*N* = 120) chose the target individuals associated with smiling context faces, against a chance level of 50% (i.e., a score of 4 out of 8). Overall children were more likely to pick the target associated with smiling faces (*M* = 5.25, *SD* = 1.84) above chance [$t(119) = 7.43$, $p < .001$, Cohen's *d* = .68].

Again following our pre-registered plan, follow-up tests were conducted to analyse performance at each age group separately. One-sample *t*-tests revealed that 5-year-olds (*N* = 40) [$t(39) = 2.83$, $p = .007$, Cohen's *d* = .45], 6-year-olds (*N* = 40) [$t(39) = 3.82$, $p < .001$, Cohen's *d* = .61], and 7-year-olds (*N* = 40) [$t(39) = 6.479$, $p < .001$, Cohen's *d* = 1.02] all chose the target associated with smiling faces above chance (see *Figure 7*).

In order to compare performance between age groups we also performed a one-way ANOVA with age as a factor. A significant effect of age was found, $F(2,117) = 3.47$, $p = .034$, $\eta^{2} = .06$. Post hoc comparisons using Tukey HSD tests indicated only one significant comparison: 5-year-olds chose the target face associated with smiling context faces significantly less than 7-year-olds ($p = .028$). No other comparisons were statistically significant (all *p*'s > .221).

In addition to our pre-registered analyses we conducted an exploratory regression to investigate further the influence of age on task performance. We regressed age in months against the number of trials on which children chose the target faces associated with happy context faces. This analysis found a significant regression equation ($F(1,118) = 6.80$, $p = .010$), with an R^2 of .055. The number of trials on which children chose the target associated with happy context faces increased by 0.53 for each month of age.

Performance on generalisation trials was assessed using separate Chi-squared tests for male and female trials. Overall, participants (*N* = 120) chose the morph of the male children shown in the happy context on 44.17% of trials. They chose the morph of the female children shown in the happy context on 53.33% of trials. The observed frequencies in the male [$X^2_{(1)}$ = 1.63, *p* = .201] and female [$X^2_{(1)}$ = .53, *p* = .465] generalisation tests did not differ significantly from chance. Further chi-squared tests revealed that this was true across all ages (all *p*'s > .114).

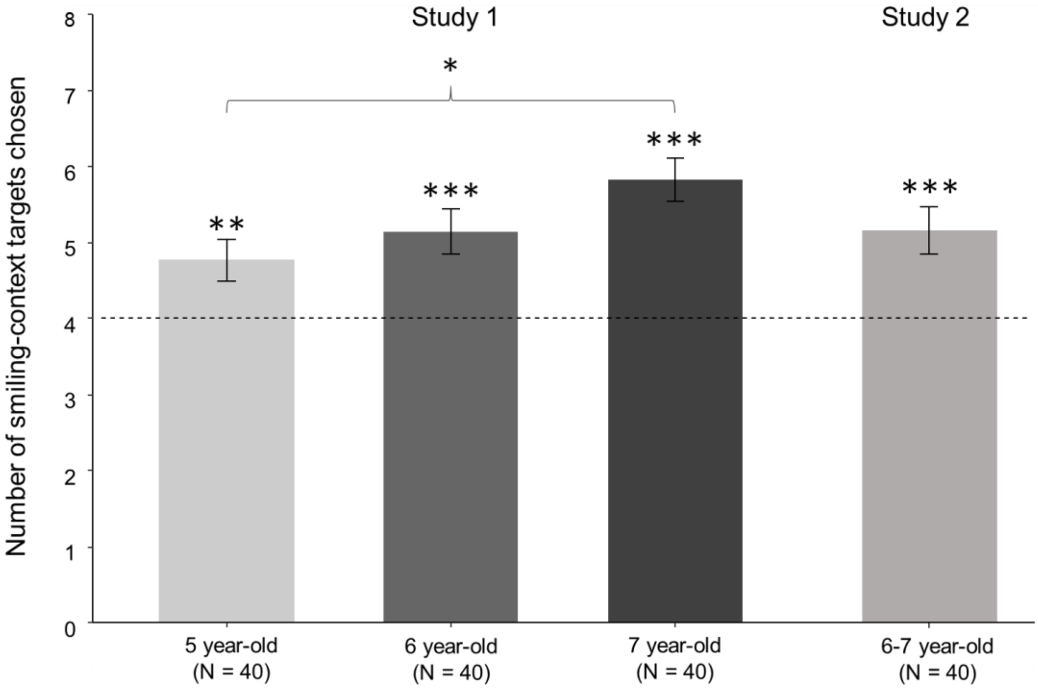

Figure 7. Results from Studies 1 & 2 by age group. Error bars represent ±SEM. Dashed line indicates chance performance. Asterisks represent a result significantly different from chance. * denotes p<.05, ** denotes p < .01; *** denotes p < .001.

Discussion

In Study 1 we showed that children aged 5-7 are able to use the non-verbal behaviour of similar aged peers to make inferences about how “nice” a target is. These results extend previous research on social referencing to the domain of faces and show that social referencing is one route through which children can learn first impressions of faces.

We did not find evidence that children generalise their learning to similar looking individuals. In order to measure generalisation, we used morphing to blend together the four target faces that had been paired with happy and fearful contexts, respectively. It was hoped that this approach would produce novel male and female identities that resembled those individuals encountered during the learning episodes. It is possible, however, that unique visual information associated with each target identity was lost. Consistent with this possibility, previous findings suggest that composite images constructed from relatively few faces start to approximate the “average” face of the population (Langlois & Roggman, 1990; Little & Hancock, 2002).

Study 2

In Study 2, we sought to further test the role of social referencing in the formation of first impressions from faces using a similar design. Similar to Study 1, there were 8 trials. Each trial was composed of 8 learning events; 4 in which one target face was seen in a smiling context, and 4 in which a second target face was presented in a fearful context. This time, however, the target faces presented at the end of each trial were morphs. Each target contained 70% of one of the two target identities and 30% of a novel identity. Thus, each target face was somewhat novel, but bore a strong resemblance to one of the trained faces. These morphs were again created using WebMorph. However, this new method of stimuli creation is better suited to the study design used here. Webmorph uses sets of facial landmarks and when creating an average, as in Study 1, the input images are combined together based on these landmarks to create an average. The consequence of this is that the output is an image that has been moved towards the average or “proto-typical” face, As a result, averaged sets of faces are often difficult to tell apart even when the input is different (Bülthoff & Zhao, 2021). In Study 2, we focus on transforming the image across one dimension (trained target identity to untrained identity) and as such preserve more of the unique facial information important for the associative processes underlying the TIM framework.

We predicted that morphs featuring target faces associated with happy context faces would be viewed as more trustworthy than morphs featuring target faces paired with fearful faces. We sought to conceptually replicate the results of Study 1 and, in doing so, test whether children would generalise from learned faces to similar looking faces.

Method

Participants

We chose to test a combined sample of 6- and 7-year-olds as children in these two age groups performed similarly in Study 1. We employed a sample size of 40 based on an *a priori* power analysis conducted using G*Power3 (Erdfelder et al., 2009). Using the overall effect size from Study 1 (Cohen’s *d* = .68) results showed that a minimum sample size of 25 participants was required to achieve power of .95. The final sample for this study consisted of 40 children (20 boys M_{age} = 83 months, age range = 72 to 94 months). An additional child was tested but excluded from analysis due to interruption during testing. Of the 40 participants included in the analysis, 39 identified as White British and 1 as English/Greek Cypriot. All children were recruited from either a science museum in an urban centre or a primary school in Northern England. Informed written consent was gathered from a parent of every child tested and assent was gained from each child. The procedure was approved by the University of York Department of Psychology's Ethics Committee and all methods were performed in accordance with the committee's guidelines.

Materials and Design

The materials and design used in Study 2 were identical to that of Study 1 with the exception of the test trials. Test trials consisted of face morphs containing 70% (shape, colour, texture) of the target faces from the training trials and 30% from previously unseen faces. Members of each pair were morphed with the same previously unseen face using WebMorph (DeBruine, 2017).

Procedure

Study 2 followed the same procedure as Study 1 with two exceptions. There was no final "generalisation trial" after test trials had finished, and children were now prompted to look at the screen before each learning event to minimise the risk of exclusions. Coding was identical to that of Study 1.

Coding

A rater naïve to the hypothesis of the study second coded 25% of the data. There was near perfect agreement between the two coder's judgements, $\kappa = .957$, $p < .001$. The one disagreement between the coders was resolved by discussion.

Results

Following our pre-registered analysis plan, we conducted a one-sample *t*-test evaluating the number of times children chose the morph featuring the target previously associated with happy context faces, against a chance level of 50% (i.e., a score of 4 out of 8). Overall children chose the target associated with smiling context faces (M = 5.18, SD = 1.99) significantly more often than chance [$t(39)$ = 3.74, p =.001, Cohen's d = .59] (see Figure 7). This replicates and extends the results of Study 1. Although we did not find evidence of generalisation in Study 1, with the more sensitive method adopted in Study 2, children did use social referencing to learn about the facial trustworthiness of individuals and generalised their social learning to very similar images.

General Discussion

We investigated how the non-verbal behaviour of others influences young children's first impressions of faces. Study 1 showed that, at least from the age of five, children use the non-verbal appraisal of similar aged peers to infer the apparent trustworthiness of others' faces. This replicates and extends previous research demonstrating that children are sensitive to the non-verbal behaviour of others, using it to inform their understanding of the social world (Skinner et al., 2017, 2019). Study 2 showed that children generalise their learning to novel individuals who clearly resemble individuals who have previously been the recipients of positive non-verbal behaviour (Richter et al., 2016; Verosky & Todorov, 2010, 2013). This supports the claim that social referencing is one route through which children can learn to form spontaneous first impressions of others' faces (Over et al., 2020; Over & Cook, 2018).

These results accord with the broader literature on social learning from nonverbal cues. Previous research has shown that young children learn intergroup biases from observing others' nonverbal responses (Castelli et al., 2012; Skinner et al., 2017, 2019) and that they use the nonverbal behaviour of teachers to infer the intelligence of their peers (Brey & Pauker, 2019). We extend this important research to the domain of first impressions from faces and show that children use the non-verbal behaviour of peers to decide which faces appear trustworthy.

Previous research has noted an apparent paradox whereby first impressions are widely shared across individuals but contain, at most, a kernel of truth (Todorov et al., 2015). Our studies suggest that social referencing is one route through which consistent but inaccurate first impressions of faces could emerge (Over & Cook, 2018). Participants in the current studies received no direct evidence relating to each target's trustworthiness. Rather, they learned about apparent trustworthiness through the nonverbal behaviour of others. While nonverbal behaviour of this sort may reflect veridical information about the targets, it can also reflect shared stereotypes (Weisbuch et al., 2009).

TIM predicts that face-trait mappings will gradually approach adult-like patterns and levels of consistency throughout development as children are exposed to more systematic messages about face-trait relationships (Over & Cook, 2018). Consistent with this view, Cogsdill and colleagues (2014) found that younger children (3- to 4-year-olds) make less consistent trait judgements than older children (7- to 10-years-old). Our finding that older children seem better able to learn about the traits of others through social referencing, reveals another interesting aspect of this developmental trajectory. Previous findings that older children exhibit stronger and more consistent trait inferences from faces may, in part, reflect older children's greater exposure to correlated face-trait mappings. However, it may also reflect the fact that older children, perhaps due to more mature face processing abilities (Taylor et al., 2004) or better categorisation skills (Short et al., 2014), are better equipped to learn about their social world.

There is broad agreement in the field that at least some first impressions from faces must be the product of learning (Germine et al., 2015; Li et al., 2020; Sutherland et al., 2020a). The work reported here builds on previous research on social referencing in other areas to demonstrate one route through this face-trait learning could occur. Importantly, however, these results do not rule out the possibility of an innate contribution to first impressions. It is possible that social learning of this type builds on an innate foundation of face-trait mappings in order to produce the consistent first impressions observed in adults (Ewing et al., 2019). It is also possible that first impressions differ in their origins and that some first impressions are more heavily reliant on social learning than others. Research in this area has shown that some first impressions are strongly influenced by the emotional expression of the target, whereas others appear to be based on the target's facial features (Olszanowski et al., 2019). It may be that social learning plays a more important role in explaining the latter than the former.

An outstanding question is whether children could generalise to more distantly related individuals. In study 2, targets closely resembled the faces used at training. Although previous research speaks to 6 & 7 year olds' ability to discriminate between highly similar faces (Gao & Maurer, 2009) it is still possible that children perceived the faces used at test as the same identities as those used in training. In future research, it would be interesting to test whether children also generalise their learning from non-verbal responses to faces that more distantly resemble the faces used at training (DeBruine, 2002; Richter et al., 2016). It would also be interesting for future research to investigate the independent effects of observing happy and fearful non-verbal displays, perhaps through incorporation of a baseline condition in which the target faces appear without context faces. Another limitation of our research is that the stimuli were entirely composed of White children. An important avenue for future research is to assess how face-trait learning generalises to more diverse stimuli and how intergroup biases may interact with face-trait learning. By focusing the study of first impressions on the developmental processes by which they are acquired, these studies suggest a number of other important avenues for future research. It would be interesting to investigate the range and limits of the inferences children make following exposure using a range of dependent variables and to explore the extent to which first impressions can be modified by altering the available cultural input. This endeavour may ultimately have applied implications as it suggests that the content of storybooks, films and TV could be manipulated in order to alter children's first impressions (Over & Cook, 2018).

Overall, these findings highlight the important role of cultural learning in explaining how children learn first impressions and, in doing so, helps explain the apparent paradox by which first impressions are widely shared between members of a community but often inaccurate (Dilger et al., 2017; Efferson & Vogt, 2013).

Chapter 4: First Impressions Through Conversation

Reference:

Eggleston, A., McCall, C., Cook, R., & Over, H. (2021). Parents reinforce the formation of first impressions in conversation with their children. *Plos one*, *16*(8), e0256118.[3]

Abstract

The tendency to form first impressions from facial appearance emerges early in development. One route through which these impressions may be learned is parent-child interaction. In Study 1, 24 parent-child dyads (children aged 5-6 years, 50% male, 83% White British) were given four computer generated faces and asked to talk about each of the characters shown. Study 2 (children aged 5-6 years, 50% male, 92% White British) followed a similar procedure using images of real faces. Across both studies, around 13% of conversation related to the perceived traits of the individuals depicted. Furthermore, parents actively reinforced their children's face-trait mappings, agreeing with the opinions they voiced on approximately 40% of occasions across both studies. Interestingly, although parents often encouraged face-trait mappings in their children, their responses to questionnaire items suggested they typically did not approve of judging others based on their appearance.

---

[3] The author, Adam Eggleston, designed the experiment, collected the data, analysed the results and wrote the article under the supervision of Dr. Harriet Over.

Introduction

Adults spontaneously attribute a wide range of traits to strangers based solely on their facial features. These first impressions include judgements about trustworthiness, honesty, competence, intelligence, aggression, and likeability (Oosterhof & Todorov, 2008; Sofer et al., 2017; Todorov et al., 2015; Zebrowitz & Montepare, 2008). While a wealth of spontaneous judgements have been studied, observers' judgments appear to load on two principal dimensions often described as 'trustworthiness' and 'dominance' (Oosterhof & Todorov, 2008; Sutherland et al., 2013). These first impressions exert a powerful influence over behaviour. For example, individuals who appear untrustworthy are less likely to be offered jobs (Olivola et al., 2014a) and more likely to face harsh sentences in criminal justice situations (Funk & Todorov, 2013; Wilson & Rule, 2015). Individuals who look competent are more likely to be elected to public office (Ballew & Todorov, 2007).

Interestingly, although some first impressions may be based on 'a kernel of truth' (Bonnefon et al., 2015; Collova et al., 2020b), many others appear unrelated to the true behavioural tendencies of the people being judged. For example, although observers show relatively high levels of agreement regarding which individuals appear trustworthy, these individuals are no more likely to act in prosocial ways than are members of the general population (Dilger et al., 2017).

Developmental research has recently begun to investigate the origins of first impressions in young children. Cogsdill and colleaugues (2014), found that children as young as 4 were able to identify which computerised faces had been manipulated to appear 'nice', 'strong' and 'smart'. Children's judgments converge with those of adults and reach adult-like levels of consistency around the age of five or six in this paradigm.

Emotional expression appears to play an important role in guiding children's reactions to others. Jessen & Grossmann (2016), found that 7-month-old infants prefer to look at faces whose features seem to resemble subtle smiles rather than subtle frowns. Later in development, children use emotional expressions to guide their behaviour: five- to 12-year-old children are more likely to invest resources in an individual who is smiling than an individual who is frowning (Ewing et al., 2015, 2019). The extent to which emotional expressions can be used to scaffold trait inferences continues to develop throughout childhood. Mondloch and colleagues (2019), found that whereas adults use emotional cues to happiness and anger in order to make judgments about likely future behaviour (e.g., "would help fight dragons" vs. "would not steal your cape"), 4- to 11-year-old children do not.

Researchers agree that learning plays an important role in the acquisition of at least some first impressions (Over & Cook, 2018). Supporting this view, research has shown that there is considerable variation in first impressions across cultures (Over et al., 2020; Shimizu et al., 2017). Further evidence comes from twin studies which demonstrate that these individual differences in trait inferences are shaped by personal experiences, rather than genes or shared environments (Sutherland et al., 2020a). Other work has shown that children form first impressions from cultural cues such as glasses (Eggleston et al., 2021). As glasses are a relatively recent product of human history, these first impressions must be learned rather than the product of an innate mechanism.

To date, relatively little research has investigated *how* first impression are learned. Recently, however, Over & Cook (2018) articulated a cultural learning perspective on the origin of first impressions. According to this view, first impressions are the result of mappings between ‘face space’ and ‘trait space’ brought about as a result of experience. Cultures consistently pair particular features of appearance with particular character traits. For example, in Western cultures villainous characters are more likely to be depicted with some kind of dermatological disorder, both in modern films (Croley et al., 2017; Ryan et al., 2018) and classic literature (Plachouri & Georgiou, 2019). Likewise, depictions of princesses in Disney films consistently pair feminine features, physical beauty, and large eyes with docility and kindness (Bazzini et al., 2010).

According to the cultural learning model, one source of face-trait mappings are social interactions between parents and children. Parents may teach their children to make judgments about other people’s characters from their physical appearance (Over et al., 2020). One route by which intergenerational transmission of face-trait mappings could occur is social referencing – children may learn how to respond to strangers that vary in physical appearance by monitoring the caregivers’ non-verbal reactions to different individuals (Fein, 1975; Over & Cook, 2018). Another route by which inter-generational transmission could occur is conversation. Parents may explicitly endorse or encourage particular face-trait mappings in conversation with their children (Over et al., 2020; Over & Cook, 2018).

Here, we investigate whether parents engage in conversations with children in which they encourage their children to make inferences about other people’s characters from their physical appearance. In Study 1, we presented children with a storybook containing images of four faces – one who appeared trustworthy (high trust face), one who appeared untrustworthy (low trust face), one who appeared competent (high competence face) and one who appeared incompetent (low competence face). We gave parents the relatively open instruction 'Talk about each of the characters shown with your child' and recorded the conversation that resulted. Of particular interest was whether parents would ever spontaneously reference trait terms such as how kind or mean the individuals in the photograph appeared and, if so, how often. We were also interested in whether parents spontaneously made reference to subtle emotional expressions of the individuals. We also wanted to explore how discussions started and how parents responded to their child's inferences, for example whether or not they reinforced the idea that the traits of individuals can be inferred from their appearance alone.

In addition to coding parents' conversations with their children, we also asked parents three questions about judging people based on their appearance. These questions related to how acceptable parents found it to judge strangers based on their appearance and how confident parents were that their first impressions were accurate. Previous research found that physiognomic beliefs, the idea that psychological characteristics can be inferred from physical facial features, are relatively common and that those who more strongly endorsed physiognomic beliefs were likely to be both overconfident in their accuracy and more reliant on physical facial cues during an economic trust game (Jaeger et al., 2020b). We were interested in the more specific question of whether or not parents' judgments would correlate with the extent to which they taught their children to judge individuals based on their appearance in a storybook paradigm.

We chose to investigate these questions with the parents of 5- and 6-year-old children. We chose this age group because we know that children in this age group appear to form some first impressions from appearance but their first impressions have not yet reached adult levels of consistency (Cogsdill et al., 2014; Mondloch et al., 2019). These studies are exploratory in nature. Rather than engaging in hypothesis testing, we sought to characterise the conversations of parents and their children on these topics. The data for all studies can be found at the OSF: (https://osf.io/3d9rf/?view_only=5710f5f555ad41c094f11f930f26e091)

Study 1

In this study, we presented parent-child dyads with a picture book containing four images. These images were of synthetic faces created using Face Gen 3.1 to appear high in trustworthiness, low in trustworthiness, high in competence and low in competence (taken from Oosterhof & Todorov (2008)). We asked parents and their children to “talk about each of the characters shown”. We measured how often parents and their children referred to the apparent traits and emotions or expressions of the individuals depicted without being explicitly prompted to do so. We also coded who initiated these conversations and how often parents reinforced the face-trait mappings of their children.

Method

Participants

Protocols were approved by the University of York’s Psychology Ethics Committee. A total of 48 individuals participated in the form of twenty-four parent-child dyads (9 Mother-Daughter, 9 Mother-Son, 3 Father-Daughter, 3 Father-Son). Participant numbers were decided in advance based on previous research exploring parent-child interactions (Chalik & Rhodes, 2015; Masur & Gleason, 1980; Rees et al., 2017; Ross et al., 2016) Of the 24 children, 12 were 5-year-olds (12 boys, Mage: 66

months, age range = 60 to 71 months) and 12 were 6-year-olds (12 boys, Mage: 77 Months, age range = 73 to 82 months). A majority of children (20/24) were described by their parents as White British. Of the remaining 4 children, 3 were described as White/Asian and 1 was described as Indian/British. All parents (Mage = 38, SDage = 8.46) confirmed that English was both their own and their child's primary language. Participants were recruited from a science museum in an urban centre where both oral and written consent was obtained, verbal assent was also elicited from children

Materials

The stimuli used in Study 1 were computer-generated face stimuli created in Face-Gen 3.1 (Oosterhof & Todorov (2008)). Stimuli were chosen based on previous research suggesting that children are sensitive to apparent variations in trustworthiness and competence in these images (Cogsdill et al., 2014). The faces were designed to be neutral on facial expression and represent high trust, low trust, high competence and low competence. The two faces used to represent each extreme of the trait were either 3 SDs above or below the average face on the particular dimension of interest (Figure 8).

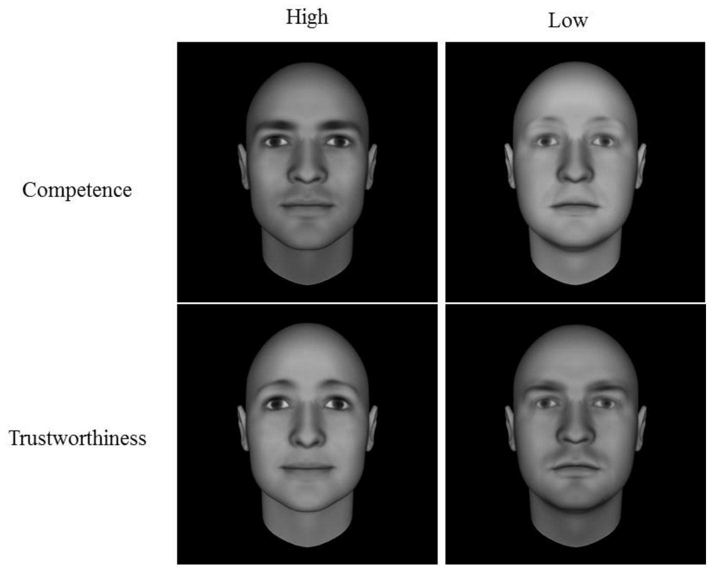

Figure 8. Computer generated face stimuli**.** All stimuli were created using Face Gen 3.1 and taken from publicly available sets of Original Computer Generated Faces, Oosterhof and Todorov (2008).

As well as the face stimuli, parents were given a three-item questionnaire to complete measuring their explicit attitudes towards judging others on first impressions. Question 1: How okay or not okay do you think it is to judge someone based only their appearance?; Question 2: How okay or not okay do you think it is to teach children to judge others based only on their appearance? Both these questions were rated on a scale of 1-7 (1 = never ok, 4 = sometimes okay, 7 = always okay. Question 3: How accurate do you think you are when forming a first impression about someone else based only on their appearance? Question 3 was rated on a scale of 1-7 (1 = never accurate, 4 = sometimes accurate, 7 = always accurate). Results for the questionnaire data can be found at the end of the results section for Study 2 as data from Studies 1 and 2 were combined.

Procedure

Participants were presented with printed versions of the stimuli formed as a book. The order of stimuli was random for each participant. The brief verbal instruction given to each parent was, “Talk about each of the characters shown with your child.” This same instruction was present on a cue card in full view throughout the experiment. Once the instruction was given the experimenter left the area to allow the participants to talk freely, these conversations were self-paced and went on for as long as the dyad desired. When finished, parents were asked to complete the 3-item questionnaire on explicit attitudes towards first impressions.

Coding

Transcription

All videos were transcribed by the first author. Transcriptions started when participants engaged with the first face and ended on participants’ last reference to the picture book. Only whole words were transcribed. From these transcriptions, four aspects of the parent/child interaction were coded for; trait terms used, amount of trait discussion, emotion/expression terms used and amount of emotion/expression discussion. After identifying trait and emotion/expression discussion we then went on to code how discussion was initiated and who initiated it, as well as parents’ responses to their child’s trait talk.

Traits

The coding scheme used the definition of a trait supplied by (Antonakis, 2011) identifying traits as individual characteristics that predict attitudes, decisions, or behaviours and consequently outcomes. Every instance of a word that fit this description was coded as a trait. Examples of trait terms used were: nice, mean, trustworthy, clever, brave and adventurous.

Trait discussion was coded as the number of words used by participants in relation to a character's traits. For instance, the below example, taken from a pair of participants, would include all words as trait discussion given that they explicitly refer to the character's trait (kind) as well as the explanation behind the label, as in Example 1.

Example 1.

Parent*:* So you think he might be, you think he might be a kind person?

Child: Yeah.

Parent: You think he might be a kind person, why do you think he might be a kind person?

Child: Because he might share toys.

Emotions

We coded references to emotional states as well as to emotional expressions. Examples of emotion terms were happy, sad, scared, smiling, tearful and frowning.

In the same way as with trait discussion, we also coded discussion about emotions. We defined this as the number of words used by participants in relation to a character's emotion including any further explanation, as in Example 2.

Example 2:

Child: Is this like a, I think, I think that's an angry face.

Parent: Is it because his mouth is like that?

Child: Yeah

Parent**:** Oh right, and what else, what else could tell you if he was angry?

Conversational initiation

We also sought to identify who initiated trait and emotion discussion, the parent or the child, and how these discussions were initiated. Initiations were coded in to one of three categories: questions (e.g. Do you think this person is nice), statements (e.g. This person is nice) or a combination of both (e.g. I think this person is nice, what do you think?).

Teaching

In order to understand whether parents teach their children face-trait mappings, we also coded whether they ever reinforced or corrected their children's trait inferences. To achieve this, we identified each time a parent responded to their child during trait discussion and coded their response in to one of four categories: reinforcement (including agreement or repetition of the child's response); correction (including rejection of child's inference or an alternative suggestion); question (including where the parent questioned the child further without endorsing their response) and other responses (including changes of subject, discussion tangential to main purpose (e.g. couldn't hear) or no follow up at all).

Second Coding

All transcriptions were coded by the first author and second coded by a rater naïve to the rationale behind the work to assess inter-rater reliability. For the purposes of second coding, transcriptions were segmented such that each time the discussion type changed to a new topic, it was labelled as a new section in the coding sheet.

These sections were then given a value of: 0 – neither trait or emotion discussion, 1 – trait discussion, 2 – emotion discussion, 3 – both trait and emotion discussion. A second coder assessed each section independently following the aforementioned coding scheme. There was near perfect agreement between the two coders' judgements, κ = .977. The few disagreements were resolved through discussion between coders.

The number of trait and emotion terms used overall by each parent-child dyad was also assessed for inter-rater reliability. There was a strong correlation between coder's judgements for traits (r=1, p<.001) and emotions (r=.992, p<.001). The few disagreements between coders were resolved through discussion between coders.

For initiation of discussion, results revealed that there was near perfect agreement between the two coders' judgements, κ = .924. Likewise, results for the inter-rater reliability analysis of parents' trait reinforcement revealed near perfect agreement between coders, κ = .971. In all cases, the few disagreements being resolved through discussion between the coders.

Results

To compare the number of words spoken by parents and children between conditions we used linear mixed models. These models included a fixed effect for condition (with the low trust condition set as the reference level) and a random effect for dyad to predict the number of words spoken. These models were fitted by restricted maximum-likelihood estimation in R (4.0.5) using the lme4 package (1.1.26). We also used the lmerTest package (3.1.3) to obtain anova tables for the fixed effects. The F and p-values from those tests are reported below. The estimates for the fixed and random effects for Study 1 can be found in Table A in Appendix B.

To test if participants were more likely to use trait words or emotion words when discussing the pictures, we used generalized linear mixed effect models to predict that binary variable (i.e., whether or not trait/emotion words were spoken at all). These models again included a fixed effect for condition and a random effect for dyad. They were fitted in R with the glmer function from the lme4 package, using a binomial (log link) as the family function. The odds ratios and random effects from these models for Study 1 are included in Tables B-C in Appendix B. For all models that revealed significant effects of condition, we used the emmeans package (1.5.5) for post hoc pairwise comparisons with a Bonferroni correction.

Preliminary analyses

On average, discussions lasted for 3 minutes 15 seconds and on average parents used 319.46 words in total during the storybook task. The linear mixed model predicting the number of words spoken by parents (see Table A in Appendix B) did not reveal a significant effect of condition ($F = 0.28$, $p = .843$). Parents used on average 79.75 (*SD* = 52.11) words while discussing the high trust face, 78.04 (*SD* = 45.54) words while discussing the low trust face, 83.17 (*SD* = 53.7) words while discussing the high competence face and 78.5 (*SD* = 56.38) words while discussing the low competence face.

On average, children spoke 130.38 words in total during the storybook task. The linear mixed model predicting the number of words spoken by children (see Table A in Appendix B) did reveal a significant effect of condition ($F = 2.99$, $p = .037$). To explore the effect of condition on children’s word count, we ran post-hoc pairwise comparisons between each condition. The only contrast to emerge as significant was between the high competence and low competence conditions whereby the low competence faces elicited more words (estimate = -12.2, t(69) = -2.82, p = .037). Children spoke 33.79 (*SD* = 31.22) words while discussing the high trust face, 34.41 (*SD* = 26.7) words while discussing the low trust face, 25 (*SD* = 15.21) words while discussing the high competence face and 37.17 (SD = 31.03) words while discussing the low competence face.

Trait terms

Topic of conversation

Overall, 13.3% of parent and child's combined conversation was about the apparent character traits of the individuals depicted. Broken down individually, traits made up 14.43% of parents' total conversation and 10.55% of children's conversation. Illustrative examples of parents' trait conversation are given below.

Example 3.

(a)

Parent*:* Has he got a friendly face or a mean face?

Child: He has, I don't know what a cross one means.

Parent: Oh, what do you think, do you think he'd be nice to you? Yeah? Okay

(b)

Parent: Do you think they're nice or do you think they're grumpy?

Child: Nice

Parent: You think they're nice. So do you think they'd be a helpful person if they came to talk to you?

(c)

Child: He looks adventurous.

Parent: He looks adventurous? Ah, that's, he does, doesn't he a bit? What else about him?

Child: He looks brave.

Parent: He looks brave? What makes you think he looks adventurous and brave?

Child: Because, the looks of his face.

Parent: The look on his face? Yeah I think I agree with you, he does look adventurous and brave doesn't he?

Child: Yeah

Other topics of conversation included references to the characters'; gender (Is this a girl or a boy do we think?), age (how old do you think he might be?), physical facial features (It's a boy, okay, and what colour eyes are his?) and occupation (What job do you reckon this man has?).

Parents

The model comparing whether or not parents used trait terms in the different conditions (see Table B in Appendix B) did not reveal any significant effects (all p's > .099). Discussion about traits made up 9.87% of parents' total conversation about the high trust face, 17.67% of parents' total conversation about the low trust face, 17.48% of parents' total conversation about the high competence face and 12.58% of parents' total conversation about the low competence face.

Use of trait terms

On average, parents used 5.71 trait terms while discussing the storybook with their children. 75% of parents used at least one trait term during the storybook task. 45.83% of parents used at least one trait term while discussing the high trust face, 58.33% of parents used at least one trait term while discussing the low trust face, 45.83% of parents used at least one trait term while discussing the high competence face and 41.67% of parents used at least one trait term while discussing the low competence face. Parents used a variety of different trait terms when describing the faces. A complete list of these trait terms, broken down by the type of face can be found in Table A in Appendix C.

The model comparing whether or not children used trait terms in the different conditions (see Table B in Appendix B) did not reveal any significant effects (all p’s > .054). Discussion about traits made up 5.55% of children’s total conversation about the high trust face, 16.59% of children’s total conversation about the low trust face, 10% of children’s total conversation about the high competence face and 9.87% of children’s total conversation about the low competence face. A visual representation of the amount of trait discussion observed for both parents and children can be found in Figure 9.

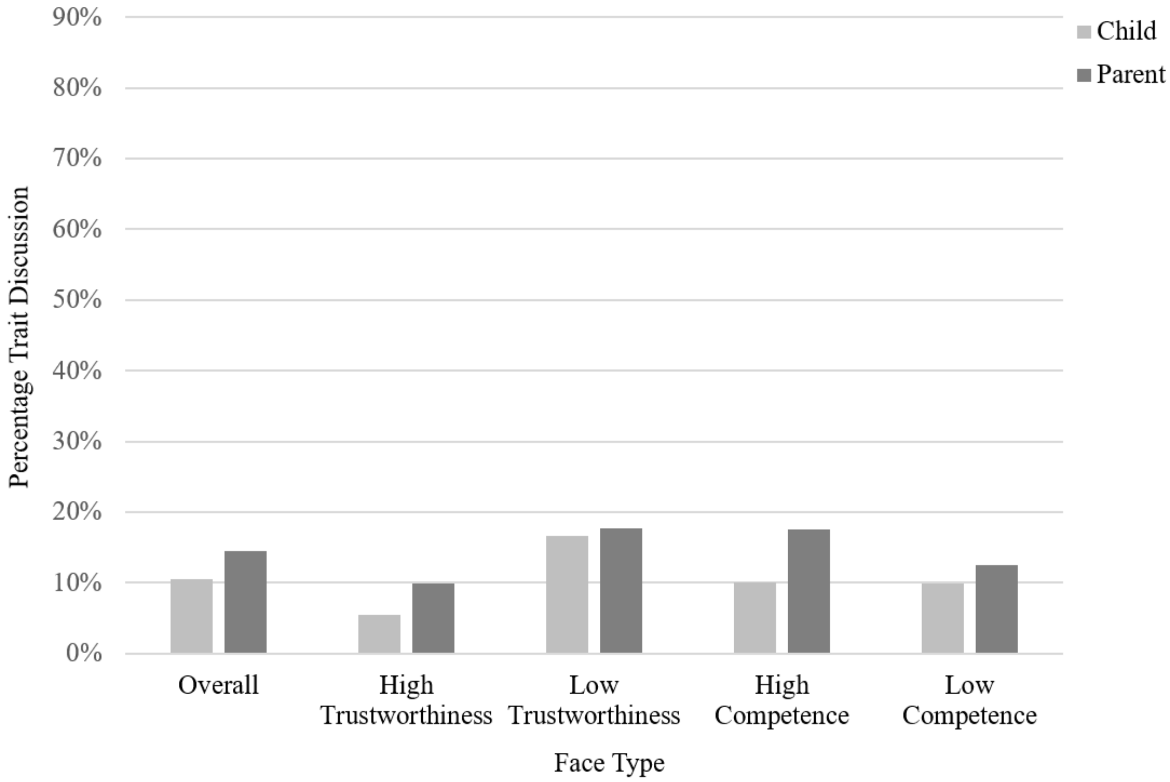

Figure 9. Study 1: Bar graph showing the percentage of parent and child conversation dedicated to trait talk, shown overall and separated by face type.

On average, children used 2.13 trait terms while discussing the storybook with their parents. A visual representation of the number of parents and children who used at least one trait term can be found in Figure 10. Overall, 66.67% of children used at least one trait term during the storybook task. 33.33% of children used at least one trait term while discussing the high trust face, 41.67% of children used at least one trait term while discussing the low trust face, 33.33% of children used at least one trait term while discussing the high competence face and 29.17% of children used at least one trait term while discussing the low competence face. Children used a variety of different trait terms when describing the faces. These terms children used broadly accord with the findings from more controlled studies (Cogsdill et al., 2014). A complete list of these trait terms, broken down by the type of face can be found in Table B in Appendix C.

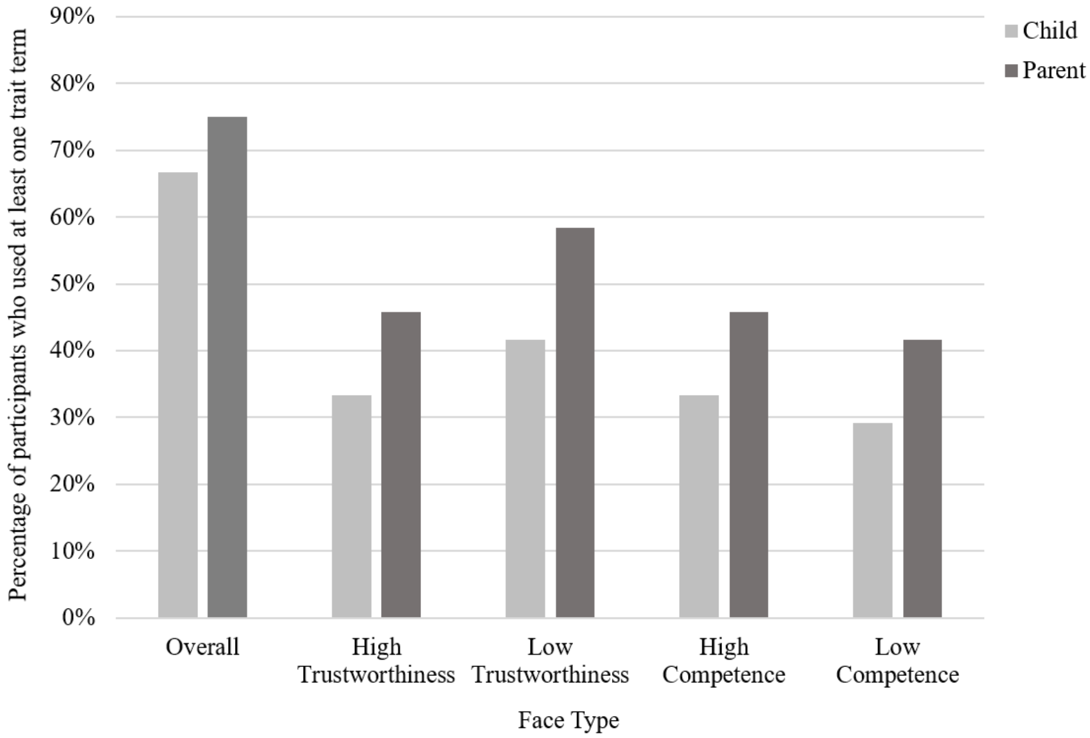

Figure 10. Study 1: Bar graph showing the percentage of parents (N=24) and children N=24) who used at least one trait term during conversation, shown overall and separated by face type.

Emotion terms

Topic of conversation.

In addition to discussing character traits, parents and children discussed the apparent emotions of each character. Overall, combined discussion about emotions made up 9.81% of parent and child’s conversation about the faces. Broken down individually, emotion discussion made up 9.82% of parents’ total conversation and 9.78% of children’s conversation. Illustrative examples of parents’ trait conversation are given below.

Example 4.

(a)

Child: Is this like a, I think, I think that’s an angry face.

Parent: Is it because his mouth is like that?

Child: Yeah.

Parent: Oh right, and what else, what else could tell you if he was angry?

Child: Red, but he isn’t red.

Parent: Oh right okay.

(b)

Parent: And are they a happy person are they a sad person?

Child: Happy that guy

Parent: ‘Cause they’ve got a smiling again is it?

Child: Yeah

(c)

Child: He looks a bit happier

Parent: Yeah, what else makes him look happy it’s not just the smile, he can, because if I smile and go like this, what else makes him look happy then?

Child: His cheeks go out wide.

Parent: Yeah, anything else? He looks like he's really happy doesn't he? Yeah.

Parents

The model comparing whether or not parents used emotion terms in the different conditions (see Table C in Appendix B) did not reveal any significant effects (all p's > .257). Overall, discussion about emotions made up 10.03% of parents' total conversation about the high trust face, 13.03% of parents' total conversation about the low trust face, 9.47% of parents' total conversation about the high competence face and 6.79% of parents' total conversation about the low competence face.

Use of emotion terms

On average, parents referred to 4.13 emotion terms while discussing the storybook with their children. 70.83% of parents used at least one emotion term during the storybook task. 41.67% of parents used at least one emotion term while discussing the high trust face, 45.83% of parents used at least one emotion term while discussing the low trust face, 41.67% of parents used at least one emotion term while discussing the high competence face and 33.33% of parents used at least one emotion term while discussing the low competence face. Parents used a variety of different emotion terms when describing the faces. A complete list of these emotion terms, broken down by the type of face, can be found in Table C in Appendix C.

Children

The model comparing whether or not children used emotion terms in the different conditions (see Table C in Appendix B) did not reveal any significant effects (all p’s > .138). Overall, discussion about emotions made up 11.96% of children’s total conversation about the high trust face, 11.26% of children’s total conversation about the low trust face, 6.33% of children’s total conversation about the high competence face and 8.74% of children’s total conversation about the low competence face.

Use of emotion terms

On average, children referred to 1.71 emotion terms while discussing the storybook with their parents. 58.33% of children used at least one emotion term during the storybook task. 41.67% of children used at least one emotion term while discussing the high trust face, 29.17% of children used at least one emotion term while discussing the low trust face, 20.83% of children used at least one emotion term while discussing the high competence face and 25% of children used at least one emotion term while discussing the low competence face. Children used a variety of different emotion terms when describing the faces. A complete list of these emotion terms, broken down by the type of face, can be found in Table D in Appendix C.

Conversational initiation

The majority of conversation about traits (73.6%) and emotions (61.4%) was initiated by parents rather than by children. Most commonly, parents introduced these topics by asking their children questions. A complete breakdown of parents’ and children’s conversational strategies can be found in Table 1.

Table 1. Study 1: Descriptive statistics for the number of times conversation was initiated by parent or child and the form of that initiation (question, statement, or a combination) for trait and emotion discussion.

	Trait (N)	Trait (%)	Emotion (N)	Emotion (%)
Total Number of Sections	72	-	57	-
Parent Initiated	**53**	**73.61%**	**35**	**61.40%**
Parent Initiated via Question	45	62.50%	28	49.12%
Parent Initiated via Statement	3	4.17%	4	7.02%
Parent Initiated via Combination	5	6.94%	3	5.26%
Child Initiated	**19**	**26.39%**	**22**	**38.60%**
Child Initiated via Question	4	5.56%	1	1.75%
Child Initiated via Statement	15	20.83%	21	36.84%
Child Initiated via Combination	0	0%	0	0%

Teaching

Parents reinforced their children’s face-trait mappings, demonstrating reinforcing behaviour on 45.05% of occasions. Parents rarely directly corrected their children’s inferences (1.1% of occasions). A breakdown of parent’s teaching behaviour can be found below in Table 2.

Table 2. Study 1: Frequency of parents' responses to child trait discussion by response type.

Response Type	Trait (N)	Trait (%)	Emotion (N)	Emotion (%)
Total Number of Sections	91	-	69	-
Reinforcement	41	45.05%	32	46.38%
Correction	1	1.10%	0	0%
Question	25	27.47%	22	31.88%
Change of subject	24	26.37%	15	21.74%

Discussion

Study 1 reveals that parents engage their children in conversations about traits inferred from purely physical characteristics. Trait conversation made up just over a 10% of overall discussion about the characters in this paradigm. This provides evidence that face-trait mappings may be formed through everyday conversations between parent and child, suggesting a wealth of opportunities for these mappings to be formed and updated. Parents often led the discussion, initiating trait discussion more frequently than did their child. Interestingly, parents often initiated these conversations using information seeking questions (Taggart et al., 2020). This suggests that parents were reinforcing the view that it is possible to draw inferences about character from appearance rather than encouraging particular inferences about the specific faces depicted.

As seen in the examples provided, children were not passive learners, they initiated some trait discussion and expressed their own trait initiation. When children made trait inferences, parents expressed their agreement with them on over 40% of occasions, suggesting that parents reinforce their children’s face-trait mappings.

We also explored conversation surrounding each character's emotional state and expression. Combined these made up over 9% of total conversation. As seen in Example 4, discussion of emotional states were often accompanied by description of the character's expression, perhaps aiding in children's emotion recognition ability which has been shown to increase significantly across the age range tested (Chronaki et al., 2015). Related to this, other work has demonstrated that 5-year-olds ability to make trait inferences such as trustworthiness vary as a function of emotional comprehension (Baccolo & Macchi Cassia, 2020) meaning that this emotion knowledge, scaffolded by parent conversation, may first be necessary before face-trait inferences can occur. Indeed, many researchers believe trait inferences to be a direct product of overgeneralisation from emotional cues (Ewing et al., 2019). Whilst the data here cannot offer causal evidence, they do point to the wealth of cultural information available to young children and one route, parent-child conversation, through which face-trait mappings could occur early in development.

Study 2

In Study 1 we demonstrate that parents engage in conversation about traits attributed to computer generated faces. In Study 2 we are interested in the same question but seek to examine conversation about images of real faces. It is possible that parents are willing to encourage first impressions about synthetic agents who don't really exist. When discussing real people, however, they might respond differently. By testing real-world faces we also hope to grant the task greater ecological validity, offering more of an insight into the types of conversations that could occur daily.

As in Study 1 we used faces that varied across the trustworthiness dimension. We also used faces that varied in perceived intelligence, akin to competence, perceived intelligence is interesting to explore given that inferences of intelligence may develop later that inferences of trustworthiness (Eggleston et al., 2021).

Again 24 parent child dyads were invited to look through a picture book containing four faces (high trust, low trust, high intelligence, low intelligence) with the instruction “talk about each of the characters shown”. Conversation was measured and is presented in the same way as Study 1.

Method

Participants

Protocols were approved by the University of York’s Psychology Ethics Committee. A total of Twenty-four parent/child dyads (7 Mother-Daughter, 6 Mother-Son, 5 Father-Daughter, 6 Father-Son) participated in the experiment. Of the 24 children, 12 were 5-year-olds (12 boys, Mage: 64 months, age range = 60 to 71 months) and 12 were 6-year-olds (12 boys, Mage: 77 Months, age range = 73 to 83 months). A majority of children (22/24) were described by their parents as White British. Of the remaining one was described as White and Black African and the other as Pakistani British. All parents (Mage = 37.96, SDage = 7.20) confirmed that English was both their own and their child’s primary language. Participants were recruited from a science museum in an urban centre where both oral and written consent was obtained, verbal assent was also elicited from children

Materials

The stimuli used in Study 2 were taken from The Karolinska Directed Emotional Faces (KDEF) (Lundqvist et al., 1998). The KDEF consists of 70 faces displaying 7 different emotional expressions. For this study only expressions previously rated as emotionally neutral were included. From the original KDEF, 66 faces had been previously rated on 14 different character traits by 327 adult participants (Oosterhof & Todorov, 2008). From these ratings those who ranked highest and lowest on judgements of trustworthiness and intelligence were selected to create 4 maximally dissimilar faces across the 2 dimensions, see Figure 11: High Intelligence (ID: AM13, Rating: 0.88); Low intelligence (ID: AM32, Rating -1.01); High Trustworthiness (ID: AM31, Rating: 1.04); Low Trustworthiness (ID: AM03, Rating: -1.56). All faces were presented in black and white. The same questionnaire reported in Study 1 was used in Study 2.

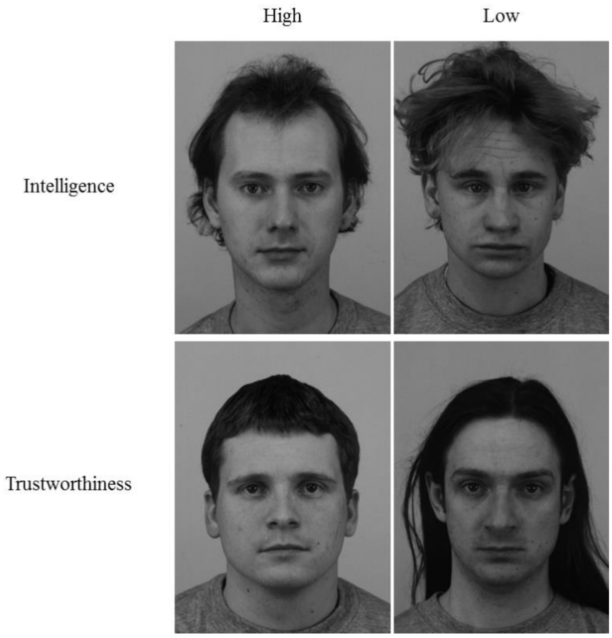

Figure 11*.* Study 2 “Real Faces”. All stimuli were taken from the Karolinska Directed Emotional Faces (Lundqvist et al., 1998). High Intelligence (ID: AM13), Low Intelligence (ID: AM32), High Trustworthiness (ID: AM31), Low Trustworthiness (ID: AMO3).

Procedure

The procedure and coding were identical to that reported in Study 1.

Second Coding

All transcriptions were coded in the same way as Study 1. As in Study 1, there was near perfect agreement between the two coders' judgements, κ = .974.There was also a strong correlation between coder's judgements for the number of trait ($r=.999$, $p<.001$) and emotion ($r=.996$, $p<.001$) terms used by each parent-child dyad.

Agreement between coders also showed strong agreement for how parents and children initiated trait and emotion discussion, κ = .838, and parent's trait reinforcement κ = .970. In all cases, the few disagreements between coders were resolved through discussion between coders.

Results

The analysis plan remained identical to Study 1. To compare the number of words spoken by parents and children between conditions we used linear mixed models. The F and p-values from those tests are included in the text. The estimates for the fixed and random effects for Study 2 can be found in Table A in Appendix D.

To test if participants were more likely to use trait words or emotion words when discussing the pictures, we used generalized linear mixed effect models to predict that binary variable (i.e., whether or not trait/emotion words were spoken at all). The odds ratios and random effects from these models for Study 2 are included in Tables B-C in Appendix D.

Preliminary analyses

On average, discussion lasted for 2 minutes 55 seconds and on average parents 300.54 words in total during the storybook task. The linear mixed model predicting the number of words spoken by parents (see Table A in Appendix D) did not reveal a significant effect of condition (F =0.72, p = .545). Parents spoke 73.17 words while discussing the high trust face, 76.63 words while discussing the low trust face, 81.42 words while discussing the high intelligence face and 69.33 words while discussing the low intelligence face.

The linear mixed model predicting the number of words spoken by children (see Table A in Appendix D) did not reveal a significant effect of condition (F =0.36, p = .782). Children spoke 23.04 words while discussing the high trust face, 25.63 words while discussing the low trust face, 23.46 words while discussing the high intelligence face and 23.21 words while discussing the low intelligence face.

Trait terms

Topic of conversation

Overall, 14.42% of parent and child's combined conversation was about was about the apparent character traits of the individuals depicted. Broken down individually, traits made up 14.36% of parents' total conversation and 14.60% of children's conversation. Illustrative examples of trait conversation are given below.

Example 5.

(a)

Parent: Does he look like a nice person or a nasty person?

Child: Nice person.

Parent: Why does he look like a nice person? 'Cause he looks like dad?

Child: Yeah

(b)

Parent: How do you think he looks?

Child: Lazy

Parent: You think he looks lazy. He looks.

Child: Grumpy, grumpy, grumpy.

Parent: You think he looks lazy and grumpy?

Child: Yeah

(c)

Parent: Do you think he looks like a good guy or a bad guy?

Child: Bad guy.

Parent: A bad guy, why do you think he looks like a bad guy?

Child: Well 'cause his face.

Parent: His face, so if you saw him in a dark alleyway would you turn around and run away?

Child: Yeah

Parents

The model comparing whether or not parents used trait terms in the different conditions (see Table B in Appendix D) found one significant effect. Here, the estimate of the "high intelligence" level of the condition factor was significant (Odds Ratio = .23, p = .047), suggesting a lower likelihood of using trait words in that condition. However, the post hoc comparisons between conditions did not reveal significant effects (all p’s > .284). Discussion about traits made up 11.33% of parents’ total conversation about the high trust face, 15.88% of parents’ total conversation about the low trust face, 12.38% of parents’ total conversation about the high intelligence face and 18.21% of parents’ total conversation about the low intelligence face.

Use of trait terms

On average, parents referred to 4.58 trait terms while discussing the storybook with their children. 75% of parents used at least one trait term during the storybook task. 41.67% of parents used at least one trait term while discussing the high trust face, 45.83% of parents used at least one trait term while discussing the low trust face, 29.17% of parents used at least one trait term while discussing the high intelligence face and 45.83% of parents used at least one trait term while discussing the low intelligence face. Parents used a variety of different trait terms when describing the faces. A complete list of these trait terms, broken down by the type of face can be found in Table A in Appendix E.

The model comparing whether or not children used trait terms in the different conditions (see Table B in Appendix D) did not reveal any significant effects (all p's > .099). Discussion about traits made up 9.95% of children's' total conversation about the high trust face, 18.37% of children's total conversation about the low trust face, 12.97% of children's total conversation about the high competence face and 16.7% of children's total conversation about the low competence face. A visual representation of the amount of trait discussion observed for both parents and children can be found in Figure 12.

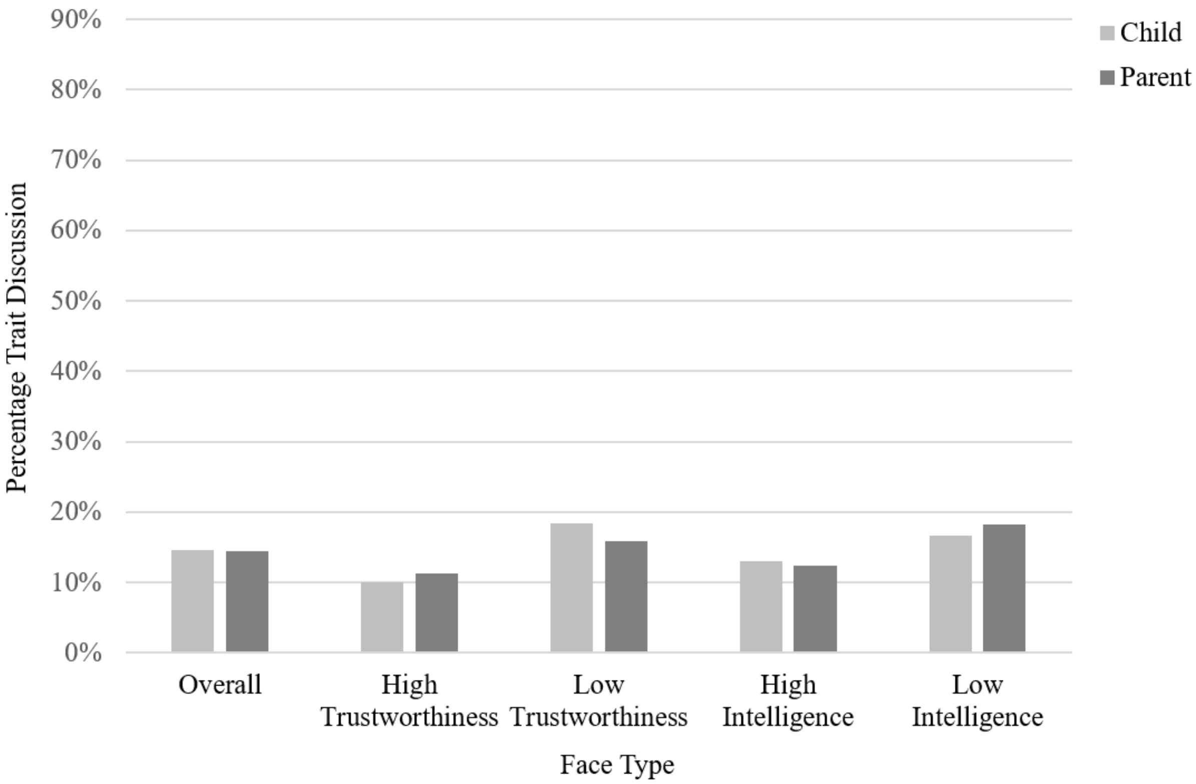

Figure 12. Study 2: Bar graph showing the percentage of parent and child conversation dedicated to trait talk, shown overall and separated by face type.

Use of trait terms

On average, children referred to 2.04 trait terms while discussing the storybook with their parents. A visual representation of the number of parents and children who used at least one trait term can be found in Figure 13. Overall 66.67% of children used at least one trait term during the storybook task. 25% of children used at least one trait term while discussing the high trust face, 45.83% of children used at least one trait term while discussing the low trust face, 20.83% of children used at least one trait term while discussing the high intelligence face and 45.83% of children used at least one trait term while discussing the low intelligence face.

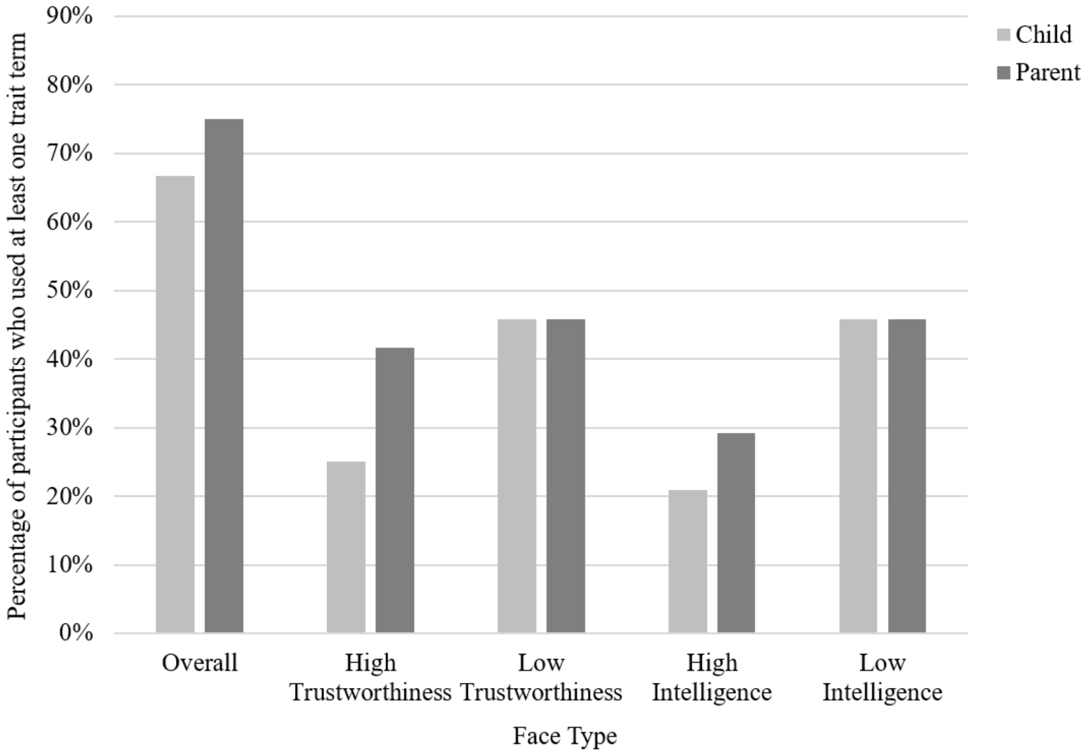

Figure 13. Study 2: Bar graph showing the percentage of parents (N=24) and children (N=24) who used at least one trait term during conversation, shown overall and separated by face type.

Children used a variety of different trait terms when describing the faces. The terms children used broadly accord with the findings from more controlled studies (e.g., 13). A complete list of these trait terms, broken down by the type of face can be found in Table B in Appendix E.

Emotion terms

Topic of conversation

In addition to discussing character traits, parents frequently discussed emotions with their children. This fits with previous research suggesting that first impressions are strongly influenced by emotional cues (Caulfield et al., 2016; Ewing et al., 2019; Montepare & Dobish, 2003).

Overall, combined discussion about emotions made up 16.42% of parent and child's conversation about the faces. Broken down individually, emotion discussion made up 16.57% of parents' total conversation and 16% of children's conversation. Illustrative examples of parents' trait conversation are given below.

Example 6.

(a)

Child: He's sad

Parent: He's sad, why do you think he's sad?

Child: Because his mouths going down.

Parent: His mouths going down, does anything else make him look sad or is it just his mouth?

Child: The mouth.

Parent: Just his mouth, okay

(b)

Child: He looks a bit sad.

Parent: He looks sad? Aww, why do you think he might be sad?

Child: Because nobody’s playing with him.

Parent: Nobody’s playing with him?

(c)

Parent: And do you think he’s happy, sad or angry or?

Child: He looks a bit sad and angry.

Parent: Sad and angry, I think so too. Because he’s not smiling is he?

Parents

The model comparing whether or not parents used emotion terms in the different conditions (see Table C in Appendix D) revealed significant effects. Here, the estimates of the “high intelligence” and “high trustworthiness” levels of the condition factor were significant (Odds Ratio = 7.62, p = .045 and Odds Ratio = 19.11, p = .010), suggesting a greater likelihood of using trait words in those conditions. However, the post hoc comparisons between conditions did not reveal significant effects (all p’s > .057). Discussion about emotions made up 19.93% of parents’ total conversation about the high trust face, 11.8% of parents’ total conversation about the low trust face, 17.5% of parents’ total conversation about the high intelligence face and 17.19% of parents’ total conversation about the low intelligence face.

Use of emotion terms

On average, parents referred to 7.46 emotion terms while discussing the storybook with their children. 83.33% of parents used at least one emotion term during the storybook task. 70.83% of parents used at least one emotion term while discussing the high trust face, 45.83% of parents used at least one emotion term while discussing the low trust face, 66.67% of parents used at least one emotion term while discussing the high intelligence face and 58.33% of parents used at least one emotion term while discussing the low intelligence face.

Parents used a variety of different emotion terms when describing the faces. A complete list of these emotion terms, broken down by the type of face, can be found in Table C in Appendix E.

Children

The model comparing whether or not children used emotion terms in the different conditions (see Table C in Appendix E) revealed a significant effect. Here, the estimate of the “high trustworthiness” level of the condition factor was significant (Odds Ratio = 23.51, p = .009), suggesting a greater likelihood of using trait words in that condition. However, again, the post hoc comparisons between conditions did not reveal significant effects (all p’s > .056). Discussion about emotions made up 19.17% of children’s’ total conversation about the high trust face, 13.5% of children’s’ total conversation about the low trust face, 17.23% of children’s’ total conversation about the high intelligence face and 14.36% of children’s’ total conversation about the low intelligence face

Use of emotion terms

On average, children referred to 2.75 emotion terms while discussing the storybook with their parents. 58.33% of children used at least one emotion term during the storybook task. 50% of children used at least one emotion term while discussing the high trust face, 37.5% of children used at least one emotion term while discussing the low trust face, 41.67% of children used at least one emotion term while discussing the high intelligence face and 37.5% of children used at least one emotion term while discussing the low intelligence face.

Children used a variety of different emotion terms when describing the faces. A complete list of these emotion terms, broken down by the type of face can be found in Table D in Appendix E.

Conversational initiation

As in Study 1, the majority of conversation about traits (56.92%) and emotions (64.94%) was initiated by parents rather than by children. Most commonly, parents introduced these topics with questions. A breakdown of how trait and emotion discussion was initiated by participants can be found below in Table 3.

Table 3. Study 2: Descriptive statistics for the number of times conversation was initiated by the parent or child and the form of the initiation (question, statement, or a combination of both) for trait and emotion discussion.

	Trait (N)	Trait (%)	Emotion (N)	Emotion (%)
Total Number of Sections	65	-	77	-
Parent Initiated	**37**	**56.92%**	**50**	**64.94%**
Parent Initiated via Question	29	44.62%	36	46.75%
Parent Initiated via Statement	3	4.62%	8	10.39%
Parent Initiated via Combination	5	7.69%	6	7.79%
Child Initiated	**28**	**43.08%**	**27**	**35.06%**
Child Initiated via Question	1	1.54%	0	0.00%
Child Initiated via Statement	26	40%	27	35.06%
Child Initiated via Combination	1	1.54%	0	0%

Teaching

Parents reinforced their child's face-trait mappings, demonstrating reinforcing behaviour on 44% of occasions. Parents rarely directly corrected their child's inferences (3% of occasions). Descriptive statistics characterising parents' responses to children's trait discussion can be found below in Table 4.

Table 4. Study 2: Frequency of parents' responses to child trait discussion by response type.

Response Type	Trait (N)	Trait (%)	Emotion (N)	Emotion (%)
Total Number of Sections	100	-	103	-
Reinforcement	44	44%	50	48.54%
Correction	3	3%	4	3.88%
Question	24	24%	34	33.01%
Change of subject	29	29%	15	14.56%

Discussion

As in Study 1 we find that over 10% of parent-child conversation centred around each character's perceived traits. Given the lack of contextual or behavioural information regarding each character, we can assume that these trait inferences are derived from each character's physical appearance. Providing at least some evidence that parents encourage face-trait mappings to be formed through everyday conversation. As in Study 1, when their children voiced trait inferences from appearance, their parents often reinforced them. Together these behaviours demonstrate a plausible route through which face-trait mappings may be formed and reinforced through everyday conversation. This extends upon the findings from

Study 1 as we presented participants with real faces, a situation more likely to reflect day-to-day reality for the parent and child.

The pattern in responses we saw in Study 1 for emotion discussion seem to be reflected in Study 2 with both parents and children describing emotional states and expressions in relation to each other. This corresponds with previous research suggesting that first impressions from appearance is closely tied to emotion understanding (Baccolo & Macchi Cassia, 2020).

Combined questionnaire data

Results

Parents' judgments about the acceptability of forming first impressions from appearance cues.

In both studies we asked parents how acceptable they found it to form first impressions of other people's characters from their appearance and how acceptable they found it to teach their children to form first impressions of other people's characters from their appearance. We combined the data from Studies 1 and 2 in order to better understand parents' answers to these questions. In general parents, judged it to be unacceptable to judge individuals based solely on their appearance. On average participants responded to the question, 'How okay or not okay do you think it is to judge someone based only their appearance?' with a mean score of 2.58 (Mode = 1, SD = 1.49). A one-sample t-test confirmed that this score was significantly lower than the possible middle score (4), $t(47) = -6.61$, $p = <.001$, $d = -0.95$. However, scores given ranged from 1 to 7, indicating that parents varied considerably in how acceptable they found judging other people on the basis of their appearance.

Parents also found it unacceptable to teach their children to judge the character of other people based on their appearance. On average participants responded to the question, 'How okay or not okay do you think it is to teach children to judge others based only on their appearance?' with a mean score of 2.56 (Mode = 1, SD = 1.61). A second one-sample t-test again confirmed that this score was significantly lower than the possible middle score (4), $t(47) = -6.19$, $p = <.001$, $d = -0.89$. Again, there was a wide variability in parents' responses with scores ranging from 1 to 7, indicating that parents varied considerably in how acceptable they found it to teach their children to treat others on the basis of their appearance. Not surprisingly, parents' answers to questions 1 and 2 were highly correlated with each other – parents who thought it acceptable to judge strangers based on appearance also thought it was okay to teach their children to do so, $r = .73$, $p < .001$.

Parents' impressions of their own accuracy in forming first impressions.

We also asked parents how accurate they felt their own first impressions were. In general, parents were not highly confident in their ability to form accurate first impressions of others' characters from their appearance. On average participants responded to the question, 'How accurate do you think you are when forming a first impression about someone else based only on their appearance?' with a mean score of 3.56 (Mode = 4, SD = 1.46). A final one-sample t-test confirmed that this score was significantly lower than the possible middle score (4), $t(47) = -2.08$, $p = .043$, $d = -0.30$. Scores given ranged from 1 to 7, indicating that parents varied considerably in how confident they were that their judgments are accurate.

Scores from questions one and two were combined to create an overall score assessing parents' belief in the acceptability of forming first impressions from appearance cues. We found a significant relationship and moderate correlation between parent’s belief in the acceptability of first impressions and their confidence that their first impressions were accurate, $r = .36$, $p = .013$.

Associations between parental attitudes and behaviour.

Interestingly, parental attitudes did not correlate with the actual extent of parental teaching about first impressions in conversation with their children. Parents overall belief in the acceptability of forming first impressions from appearance cues did not correlate with either the number of trait terms parents used in conversation with their children nor the percentage of words used to discuss a character’s traits, (all *p*s>.773). Likewise, parents’ confidence in their own first impressions did not correlate with their use of trait terms nor the percentage of words used to discuss a character’s traits (all *p*s>.505). Although these results must be interpreted with considerable caution due to the modest sample size, they suggest that there is not a strong relationship between parents’ explicit attitudes about the acceptability of judging people on appearance and their actual tendency to teach associations between appearance and character to their children.

Discussion

Questionnaire data revealed that parents generally think it is unacceptable to judge others based off their physical appearance. However, responses revealed that opinion varied widely when considering whether forming impressions from appearance is an acceptable and worthy pedagogical goal. In line with previous research, those who did endorse judging others on their appearance were also more confident that their first impressions were accurate (Jaeger et al., 2020b).

Comparing parents' questionnaire responses to their task performance revealed that these explicit opinions did not influence their actual interactions, at least in this paradigm. Interestingly, parents who refused to endorse judgments based on first impressions were just as likely to engage in conversation about traits based purely on physical features.

General Discussion

Across two studies we aimed to investigate the important question of how first impressions may be learned. Previous research adopting a cultural learning perspective has suggested one possible way through which the inter-generational transmission of face-trait mappings could occur is through parent led conversation. In support of this, these data seem to show that parents do sometimes engage their children in conversation about the character traits attributed to unfamiliar individuals on the basis of their physical appearance. Parents engaged in this type of conversation both when discussing computer generated faces (Study 1) and real world faces (Study 2). In line with our assumption that face-trait mappings are facilitated through parent-led conversation we found that, across both studies, parents tended to initiate these conversations, often encouraging their child to make trait inferences through the use of questions. Interestingly for our purposes, parents did this even though no explicit instruction was given to talk about the personalities of the individuals depicted. Taken together, these studies suggest that children are regularly exposed to social situations that could plausibly play a role in teaching them that it is possible to judge others' character traits from their appearance (Over et al., 2020; Over & Cook, 2018).

Further analysis of our data suggest that parents explicitly teach their children face-trait mappings, reinforcing the inferences children make approximately 40% of the time across both studies. These data suggest that, at least by the age of five, children have substantial opportunities to socially learn the face-trait mappings common within their culture. It is plausible that parental teaching is one mechanism through which children learn first impressions that are common within their culture even when they lack validity – i.e., they do not reflect the actual character traits of the individuals being judged (Over & Cook, 2018).

It is interesting to note that children were active participants in the conversations we recorded, commenting on the apparent character traits of the individuals depicted themselves. This accords with previous research suggesting that, at least by the age of five, children form consistent first impressions of others (Cogsdill et al., 2014; Cogsdill & Banaji, 2015). In future research, it would be interesting to investigate whether parents talk to even younger children about the apparent character traits of novel individuals and to examine in what ways parental conversations with their children change over time. Studies with younger children would help disambiguate whether parents create face-trait mappings in their children as well as reinforcing the face-trait mappings their children already possess.

While these data highlight the wealth of social information available to children regarding how appearance relates to character, they do not provide evidence that these types of social experiences play a causal role in children's developing first impressions. It would be interesting to manipulate how an experimenter talks to children about faces and then measure whether this influences children's first impressions on a judgment task. The types of parental conversation recorded in this study could provide a useful starting point for developing such a manipulation.

Of further interest is the finding that, in both studies, parents and children spoke about the emotions of the individuals depicted as well as their apparent character traits. This was the case even though participants had been given no prompting to do so and the stimuli used in both studies were designed to be emotionally neutral (Lundqvist et al., 1998; Oosterhof & Todorov, 2008). These findings underscore previous research showing a tight connection between emotional expression and trait judgments (Tang et al., 2019; Willis et al., 2011). This connection may be explored in future work quantifying the overlap in emotional and trait related conversation which is instead described in isolation from each other here. One possible explanation for this connection is offered by the 'emotional overgeneralisation hypothesis'. According to this hypothesis, individuals whose facial features subtly resemble smiles tend to be judged more trustworthy than individuals whose facial features subtle resemble frowns (Said et al., 2009; Zebrowitz, 2004). An alternative explanation is that the extent to which faces used in first impressions research are truly 'emotionally neutral' may have been overestimated in previous work. Developmental work that seeks to investigate first impressions in the absence of emotional cues may wish to control their stimuli more closely (Jessen & Grossmann, 2016; Sakuta et al., 2018).

A further interesting aspect of our findings relates to parents' explicit rejection of judging others based on their appearance. Even though all parents and their children engaged in at least some discussion about the apparent character traits of the individuals depicted, parents tended to state that it was inappropriate to judge others based on their appearance. These results must be interpreted with caution because of the modest sample size and the lack of anonymity in parents' responses. Future research may consider collecting larger samples in more anonymous settings, for example, through online data collection. Nevertheless, our data do suggest some interesting possibilities for further work. If future work seeks to modify the ways in which parents teach their children about first impressions, our research suggests it will be important to develop interventions that target their actual teaching behaviour rather than merely their attitudes about teaching,

Previous research has shown that trait judgments emerge early in development (Charlesworth et al., 2019; Cogsdill et al., 2014; Cogsdill & Banaji, 2015) and suggested that learning plays a role in the acquisition of these judgments (Eggleston et al., 2021; Over & Cook, 2018). The research reported here moves beyond previous research by starting to investigate *how* this learning takes place. In doing so, it opens up a number of interesting avenues for research on first impressions. For example, it will be important to explore how conversations between parents and children differ depending on the nature of the faces depicted. In these studies, we presented parents and children with picture books containing images of White individuals. In future research, it will be important to vary the ethnicity of the individuals depicted. Exploring how parental conversation varies depending on the group membership of the individuals depicted would help integrate the study of first impressions with research on stereotyping and prejudice. Previous research has shown that White parents are often reluctant to discuss ‘race’ and racism with their children (Zucker & Patterson, 2018). In this context, it would be very interesting to determine whether trait discussion could capture implicit biases in parental conversation.

Whereas we chose to focus on verbal behaviour to understand parent-child interaction, it will be important for future research to investigate how the non-verbal behaviour displayed by parents influences children’s inferences about traits. Non-verbal behaviour such as emotional expression and gesturing have been shown to impact children’s social judgements (Over & Cook, 2018; Skinner et al., 2019). Future research could analyse the valence of parents’ initial expression when each face is revealed and how it varies according to the particular face depicted.

Future work may also investigate parent-child conversation in the absence of any instruction to talk about the faces depicted. One route by which to do this would be to give parents and children a seemingly unrelated task, such as memorising the faces, and measuring incidental conversation about traits. Another route by which to achieve this would to be to analyse corpus data for evidence of naturally occurring conversations about face-trait mappings.

Finally, it would be interesting to investigate how the composition and cultural background of parent-child dyads influences conversations about the apparent traits and emotions of the individuals depicted. Previous research suggests that mothers may be more likely to make references to emotions, and to use causal explanatory language when referring to emotions, than fathers (LaBounty et al., 2008). Furthermore, there are systematic cultural differences in the first impressions that individuals form which may manifest themselves in different styles of parent-child interaction (Han et al., 2018; Wang et al., 2019).

The study of first impressions is becoming increasingly prominent within the developmental literature. Recent research has investigated the developmental origins (Jessen & Grossmann, 2016, 2017) and behavioural consequences (Ewing et al., 2015) of first impressions among children. We contributed to this work by exploring one of the developmental mechanisms through which first impressions may be acquired and/or reinforced. Our data suggest that parental conversation is one plausible mechanism through which first impressions could be learned (Over & Cook, 2018).

Chapter 5: General Thesis Discussion

The findings presented from each empirical chapter make important contributions to work investigating the origins of first impressions. In Chapter 2, I investigated whether first impressions could plausibly be learned. Adults showed sensitivity to glasses as a cue to intelligence (Study 1), with this sensitivity remaining after minimal exposure (Study 2) and instruction to ignore the cue (Study 3). The pattern of responses of adults were very similar across all three studies reflecting the speed and automatic nature of first impressions from cultural cues. Taken together, these studies suggest that the speed and automaticity of trait judgments are not uniquely supportive of a nativist view of their origins. The developmental origins of first impressions were also explored in this chapter, revealing that 6-year-old, but not 4-year-old, children were sensitive to both cultural and physical cues to intelligence. The developmental pattern of responses were broadly similar for cultural and physical trials, suggesting that the same underlying cognitive mechanism may underlie impressions from both types of cue.

Chapters 3 and 4 investigated how first impressions could be learned, focusing in particular on the role of social learning. The results from Chapter 3 reveal one route through which first impressions could emerge. Across both studies, children aged 5-7 years were sensitive to the non-verbal behaviour of their peers, using this information to make inferences about a target's "niceness". Study 2 went on to demonstrate the ability for children to generalise the learned first impression to include novel targets who resemble the learned target.

Chapter 4 explored another means by which children could learn the first impressions prevalent within their culture. In particular, it investigated how parent-child interactions may serve the development and maintenance of first impressions. Across two studies, adult-child interactions were analysed during a storybook task. Despite no explicit instruction to do so, the apparent traits of the characters in each book were discussed. Parents were both more likely to initiate these discussions than their child was and to be more likely to reinforce their child's trait inferences rather than correct or dismiss them. Across both studies around 13% of overall discussion focused on the potential traits of each character.

Throughout this thesis I have presented arguments in support of social learning theories, such as the TIM framework, and how these theories offer an explanation for how all first impressions may be acquired via learned face-trait mappings. However, this is not to say that first impressions are the sole result of learning. Many researchers agree that first impressions are the product of a mix of learning and innate mechanisms (Ewing et al., 2019; Over & Cook, 2018; Schaller, 2008; Siddique et al, 2022). Indeed, even the ability to form links between specific aspects of facial structure and personality traits relies on some type of innate architecture for associative processes to take place. The domain-general associative mechanisms that mediate this learning undoubtedly evolved genetically as a means of tracking predictive relationships and are at the root of many social behaviours (Heyes & Pearce, 2015). Furthermore, TIM describes how other innate mechanisms may help to facilitate social learning opportunities and so first impressions (Over & Cook 2018) such as a high interest in faces from birth (Johnson, 2005; Johnson et al., 1991). Where, however, current social learning and evolutionary theories seem to diverge is on the theory that domain-specific mechanisms, because of adaptive pressures, are responsible for first impressions. This pairs with the idea that there has to be some sort of innate understanding of other’s traits in relation to a particular face percept. As demonstrated throughout this thesis, common evidence used to support such evolutionary theories seems to be equally compatible with social learning accounts. Without clarity from researchers on what particular aspect of first impressions are innately specified it is difficult to detail a so called “hybrid model” (Sutherland et al., 2020c) of first impressions which considers both accounts.

Overall, the research presented in this thesis speaks to the plausibility of cultural learning accounts for first impressions. Identifying the early sensitivity to, and automatic processing of, culturally learned cues to first impressions as well as beginning to explore possible learning routes. The empirical work presented in this thesis represents the early stages of developmental work on first impressions. As this area of the first impression literature continues to grow, there are certain aspects that warrant further investigation and outstanding questions still to be addressed. Both of which will be considered in the remainder of this section.

Directions and Suggestions for Future Research

Sources of First Impressions

Parents & Peers

The empirical chapters presented in this thesis begin to illuminate how important social learning may be in the development and inter-generational transmission of first impressions. Two potential sources of first impressions, via social learning, explored are that of a child’s peer group and their parents. One important aspect of social learning to consider is whether the information garnered from both types of informant, peers and parents, are weighted the same by children and how their influence might change throughout development. Previous research has demonstrated that as children age, their learning preferences shift from their caregiver to strangers, experts, and peers (Henrich & Broesch, 2011; Lucas et al., 2017; Reyes-García et al., 2016). This previous work is relevant to the studies presented in Chapter 3 whereby children showed their sensitivity to the non-verbal behaviour of their peers when forming impressions. The question raised is whether this sensitivity would be the same or different when learning from adults or parents, and whether this changes across development. An interesting way to test the relative influence of parents and peers would be for children to view contrasting non-verbal behavioural responses of parents and peers to target faces (e.g., parent is smiling, peer is fearful) and measure if there is a developmental change in children's use of peer over parental endorsement when making trait inferences. This would concur with previous research highlighting the influence of similar aged peers on children's learning preferences (DeJesus et al., 2018; Hennefield & Markson, 2017; Tasimi et al., 2015), with children's strategy moving towards learning from a peer in cases where an aged-matched source may be more knowledgeable (Vanderborght & Jaswal, 2009), for example in Chapter 3, how nice a peer might be. A secondary question could also be whether the influence of an adult or peer on first impressions could vary depending on the trait in question. For example, a child may value the opinion of an adult (e.g. a teacher) more highly then that of a peer when it comes to evaluating a person's competence, but value a peer when evaluating a person's niceness.

Cross-cultural approaches

Developmental work is particularly useful in its ability to inform researchers on the early acquisition of first impressions including the sources, and number, of learning opportunities open to children. However, more work needs to be done in recognising the biases present in current research and how they limit the explanatory power of work on first impressions. For example, research on first impressions, including the work presented here, often focuses on using White faces as stimuli (Cook & Over, 2021). Related to this problem is the fact that developmental research as a whole often relies on data collected from WEIRD (Western, educated, industrialized, rich, and democratic) populations with few cross-cultural comparisons (Amir & McAuliffe, 2020; Nielsen et al., 2017; Nielsen & Haun, 2016). The consequences of this lack in diversity of both stimuli and participants is that the universality and accuracy of first impressions in current work may be over-stated (Cook & Over, 2021). This is especially important when considering the extent to which first impressions may have an innate basis or are culturally learned. The studies presented in Chapter 2 speak to the plausibility of socially learned first impressions, demonstrating the link between modern cues such as glasses and the trait of intelligence. Social learning theories such as TIM can explain mappings between modern cues and traits but also suggest that cultural differences in these mappings will be predictable based on observed cultural input (Over & Cook, 2018). This possibility can only be explored however through diversifying both stimuli and samples. Culture-specific rituals may be one source of first impressions that results in predictable and measurable differences in trait associations (Over et al., 2020), and so a fruitful area for future developmental research to focus on.

Media

Media images have consistently been shown to be sources of stereotyped information, with consistent and predictable stereotypical character depictions observable in TV programmes (Dobrow & Gidney, 1998; McDade-Montez et al., 2017; Northup & Liebler, 2010), video games (Dill & Thill, 2007; Gestos et al., 2018; Miller & Summers, 2007) and children's books (Diekman & Murnen, 2004; Hollis-Sawyer & Cuevas, 2013; Lewis et al., 2020). As noted in Chapter 3, exposure to racially biased non-verbal cues on TV and video have been shown to influence racial attitudes (Castelli et al., 2012; Weisbuch et al., 2009; Willard et al., 2015). Whilst a majority of the work cited thus far has focused on consistent messaging regarding a person's traits depending on their gender and race, recent work has also demonstrated frequent associations between depictions of physical facial characteristics (e.g., dermatological conditions) and traits (e.g., evil character) in film (Croley et al., 2017; Ryan et al., 2018) and literature (Plachouri & Georgiou, 2019). It may be that media depicting consistent face-trait associations is a key medium through which wide-spread first impressions are learned and diffused within a culture (Over & Cook, 2018). The extent to which digital media can influence wide-spread first impressions still needs to be explored.

Given the rise in digital media use also seen in young children (Rideout & Robb, 2020); the question of digital media's influence on first impressions may be particularly relevant to developmental research. Recent work suggests that the increase in children's screen time presents more opportunities for social learning than any other time in human history, and so children may now be particularly susceptible to digital learning (Nielsen et al., 2021; Strouse & Samson, 2021). Recent experimental work has demonstrated this through children's increased trust in an online source or teacher over that of a peer (Wang et al., 2019) and through young children's increased normative behaviour when a task was demonstrated on a laptop screen compared to a live presentation (Fong et al., 2021). Due to the current global Covid-19 pandemic and mask mandates, children's opportunities for real life face-face interactions may be more limited. The consequences of this have already been describe in work demonstrating children (and adults) reduced face perception and emotion recognition abilities (Stajduhar et al., 2022, Tsantani et al., 2022). The reduced opportunities to form face-trait mappings from everyday social interactions may mean that there is an increased reliance on learning from media depictions. Previously, access to consistent media depictions has been described as possibly increasing inter-rater agreement (Zebrowitz et al,, 2012) and this possibility will be interesting to consider and track across development.

Pre-requisites of First Impressions

What general cognitive abilities and experiences might first impressions depend on? This question is especially important when considering the developmental differences in the sensitivity to cues used in impression formation, seen in Chapters 2 & 3, as well as within other developmental first impression literature (Charlesworth et al., 2019; Cogsdill et al., 2014; Cogsdill & Banaji, 2015; Mondloch et al., 2019; Palmquist et al., 2019; Palmquist & DeAngelis, 2020).

Expertise Accounts

A major role in the acquisition of first impressions is likely to be experience (Over & Cook, 2018), shown to be an important factor in explaining both first impressions and idiosyncratic differences in these impressions (Hehman et al., 2017; Sutherland et al., 2020a; Xie et al., 2019). Within each chapter of this thesis the plausibility of the TIM framework has been discussed, describing first impressions as learned associations between separate regions in the brain, face space and trait space (Over & Cook, 2018). Expertise accounts in the past have largely focused on face space and its development. In adults, advanced face processing abilities have been associated with the use of norm-based coding whereby faces are encoded in relation to a face prototype (Valentine, 2018). This prototype serves as an average of all faces previously encountered, with different facial dimensions (e.g. eye size) represented by a unique vector in face-space. New faces are compared against this prototype and the face prototype itself is continuously being updated. In line with this, research suggests that early childhood experiences help children to categorise people based on their facial characteristics, via norm-based coding (Short et al., 2014). Younger children are more likely to have refined spaces for social groups that they have had more exposure to, such as own race faces (Short et al., 2011) with specific prototypes (including age, sex and race) becoming more refined and occurring later due to experience (Short et al., 2014; Zhou et al., 2018). In terms of first impressions, this could mean that a certain level of experience is necessary for a trait specific prototype to emerge and so for increasingly complex first impressions to develop. This is in line with the theory that refinement of face-trait mappings is dependent on environmental input and experience (Over & Cook, 2018)

Some work speaks to experience based accounts for general face processing, demonstrating that extroverts, who are more likely to engage in face-face social interactions, have increased face processing abilities compared to introverts due to increased levels of experience (Baccolo & Macchi Cassia, 2019). Recent developmental work supports theories of expertise in direct relation to first impressions and trait inferences, demonstrating increased levels of differentiation between trustworthy faces with age across 5-year-olds, 7-year-olds and adults (Baccolo & Macchi Cassia, 2020).

Emotion Recognition

In the first impression literature, the association between emotional expression and judgements of trustworthiness has been shown in children as young as 5 years (Caulfield et al., 2016; Tang et al., 2019). Explored partly in Chapters 3 & 4, the link between emotion and first impressions is well documented (Kocsor et al., 2019; Said et al., 2009; Willis et al., 2011), indicating emotional expression as an important driving force behind first impressions. Developmental work on this link describes children's trusting behaviour based on emotional expression as adult like from an early age (Ewing et al., 2015, 2019). Ewing and Colleagues found that across age groups (5 years, 10 years, adult) the same pattern of trusting behaviour could be observed during an economic trust game, with investments greatest for happy faces. This consistency in behaviour across age groups has been used as an argument that emotional expression is an especially important cue informing early developing first impressions (Ewing et al., 2019). A sentiment somewhat supported by recent findings linking children's emotion comprehension skills and first impressions. Baccolo & Cassia (2020), found that 5-year-old's, but not 7-year-old's, scores on an emotion-understanding test was positively correlated with their performance on a trustworthiness discrimination task. Responses to emotional expressions, however, are not universal (Gendron et al., 2014), with some cultures associating smiling faces with overall positive qualities and others associating them with negative traits such as low intelligence and untrustworthy (Krys et al., 2014, 2016). Therefore, it would be interesting to assess whether these results replicate cross-culturally and whether or not sensitivity to emotion cues to traits reported early on in development (e.g. happiness = trustworthy) are initially universal and change slowly across development.

Some developmental work calls in to question the importance of emotional expression on early first impressions altogether, with consistent use of emotional expression to inform first impressions not occurring until late childhood (Caulfield et al., 2016; Mondloch et al., 2019; Van Der Zant et al., 2021). For example, Mondloch and colleagues (2019) found that for both subtle and explicit displays of emotion children, 4-11 years-old, did not display adult-like patterns of behaviour. Reporting that, in situations requiring either a trustworthy or a dominant partner, children's selection did not reveal a combined influence of facial expression and target trait. These latter findings contrast with previous work suggesting a strong role of emotional expression in trait judgements (see Chapter 1, overgeneralisation hypothesis). They suggest that whilst children around 4 years can make appropriate trait judgements to inform their first impressions (Liu et al., 2007; Mondloch et al., 2019), these judgements may not rely heavily on sensitivity to subtle forms of emotional expression (Mondloch et al., 2019; van der Zant et al., 2021). A pattern seen more so in adults (Kocsor et al., 2019; Said et al., 2009; Willis et al., 2011). As recent work has identified that the dimensions on which children's faces are judged (shyness and niceness) may be different to that of adults (trustworthiness and dominance) (Collova et al., 2019). Work should continue to assess the strength of the relationship between emotion and first impressions across these new dimensions, which may be of particular importance in assessing children's level of peer acceptance. For example, early sensitivity to fear cues on impressions of shyness, and how this relates to children's choice of play partners, could be investigated (van der Zant et al., 2021).

Multiple Cues to First Impressions

This thesis, and a majority of work in the first impression literature, focuses on impressions driven by facial characteristics. Whilst this has been an extremely fruitful area of research with the capacity to explain a range of behavioural biases, including in sentencing (Wilson & Rule, 2015) and voting decisions (Olivola & Todorov, 2010), limiting first impressions to faces alone may be too reductive. As already demonstrated in Chapter 2, failing to account for the impact of a diverse range of cues to first impressions (e.g. cultural cues) can result in an underdeveloped understanding of the phenomena and the mechanisms that underlie it. In this section, the influence of multiple different types of cue will be discussed.

Body

Recent work has attempted to expand upon research exploring first impressions from emotional expression by also considering impressions from expressive behaviour. In a study pitting these two types of cue against each other Van Der Zant et al (2021) found that, for adults and children, emotion expressed through posture was more important when judging a person’s dominance but facial features were key to trustworthiness impressions. The importance of bodily cues have also been identified in other work suggesting that reliable (but inaccurate) first impressions can occur from viewing a target’s walking gait (Blaskovits & Bennell, 2019; Thoresen et al., 2012) and that judgements of facial expression are influenced by posture (Nelson & Mondloch, 2017). Altogether, this work implies that rather than a specific mechanism associated with overgeneralisations from emotional facial cues, first impressions may occur as a result of inferences from multiple cue types. Furthermore, extracting information from multiple cues, rather than just those from the face, may even increase their accuracy (Rule et al., 2012), highlighting the need for researchers studying first impressions to consider the role of more than just the face on first impressions. Recent work investigating children's accuracy of peer's niceness and shyness revealed modest accuracy for judgements of niceness, but not shyness (Collova et al., 2020a). Based on the work reviewed in this section, future studies may benefit from investigating the role of multiple cues, such as child posture, which may be particularly important to perceptions of shyness, increasing levels of accuracy.

Vocal

One key finding to take from Van Der Zant and colleague's (2021) work is that the salience of a particular cue may differ depending on the trait being inferred. In line with this is research investigating the role of a target's voice on the impression they emit, with voices appearing to have larger and more reliable effects for dominance, but not trustworthiness judgements, compared to faces (Mileva et al., 2017; Rezlescu et al., 2015). Impressions from vocal cues seem to be automatic (Mileva et al., 2017), similar to the impressions from faces described in Chapter 2. More parallels can be drawn between the use of facial and vocal cues in that first impressions from voices are reliable (McAleer et al., 2014), can influence participant's voting choices (Tigue et al., 2012) and influence perception of a target's leadership ability (Klofstad et al., 2012). However, again as with faces, the usefulness of impressions drawn from the voice may be limited due to the high within-person variability of judgements from the same target (Lavan et al., 2020, 2021) and low accuracy of trustworthiness based impressions (Rezlescu et al., 2015; Schild et al., 2020). Furthermore, the power of voices to influence behaviour may be lower than that of faces, given that the trustworthiness of voices did not influence behaviour in an economic trust game (Knight et al., 2021) whereas trustworthy faces have repeatedly been shown to (Chang et al., 2010; Ewing et al., 2015; Rezlescu et al., 2012).

Given this low accuracy, and perhaps small influence on behaviour, the same question appears as for first impressions from faces, from where might these impressions originate? This is for future developmental work to explore in more detail. Current work, however, hints at experience-based explanations, with infants' preference towards those with native accents emerging by 5 months of age (Kinzler et al., 2007). Given how much human language varies this preference is plausibly learned. In line with TIM, media could also play a role in learning mappings between particular vocal characteristics and spaces in trait space. Analysis of children's media has frequently shown the use of language/voice to convey some aspect of a character's personality (Dobrow & Gidney, 1998; Rosa, 2017). For example, in US programming, villains are often portrayed with a foreign accent (Dobrow & Gidney, 1998) and the use of stereotyped accents in children's media more generally has been described as "a vehicle by which children learn to associate specific characteristics and lifestyle with specific social groups." (Lippi-Green, 2012).

Cultural Artifacts

Similar to vocal and body cues, clothing can influence first impressions and subsequent behaviour. Depending on the clothing being worn a person may be judged as more (or less) intelligent (Behling & Williams, 2016; Morris et al., 1996), charismatic (Maran et al., 2021), and aggressive (Vrij, 1997). What is unique about clothing compared to the other cues discussed so far is that it is fully changeable, giving a person much more control over the impression that someone might form of them. This is sometimes used a social tool, allowing a person to quickly convey their membership to a particular social group. For example, the colour tie a politician wears can signal their party alignment. Children as young as 4 years already start to show sensitivity to clothing cues, identifying formally dressed targets as more knowledgeable than a casually dressed one (McDonald & Ma, 2015). Work linking the fields of impressions from facial features and clothing has begun to explore their combined influence in adults. Oh, Shafir & Todorov (2020) found that after a brief exposure of 129ms clothing had a significant effect on facial competence judgements. Other work has noted the interaction between clothing and postural cues to first impressions, with different competence ratings occurring depending on both a person's attire and posture (Gurney et al., 2017). Combined, this work again highlights the importance of considering the interaction of multiple cues to first impressions.

The way someone is dressed introduces a whole host of cues that a person can use to form an impression of a target. These modern artefacts, not rooted in biology, are interesting to consider when it comes to the origins of first impressions, as any associations with traits must be learned. As already demonstrated in Chapter 2, there is early sensitivity to these cultural cues (McDonald & Ma, 2015). A challenge for strong nativist accounts is that, stating impressions from faces are governed by an innately specified mechanism, implies that the impressions formed from modern cultural artefacts are arrived at differently. A more parsimonious view is that first impressions from faces, posture, clothing etc. are acquired through learned associations between a representation of the particular cue and a point in a unified trait space (TIM). The ability to form such mappings early on in development could be further explored through the acquisition of first impressions following brief training

(Lee et al., 2021). In a series of experiments with adults, Lee and colleagues (2021) were able to train participants to associate the physical characteristics of a novel set of stimuli (Greebles) with trustworthiness and competence. Judgements across these two dimensions were dependent on multiple cues and generalised to never before seen stimuli exhibiting similar physical characteristics. The replication of this work with children would demonstrate the ease with which trait mappings may be learned and generalised to novel targets.

Interventions

Given the consequences of first impressions from faces, as well as other cues, exploration into possible interventions designed to mitigate their deleterious effects seems warranted. If first impressions are the product of a social learning mechanism then the implication is that these impressions may be malleable and subject to substantial change. This section will explore what current research reports regarding the potential effectiveness of different intervention approaches.

Targeted Interventions

The impact of interventions on first impressions from faces is currently not widely explored. The few studies that do exist focus on how targeted interventions can help to increase positive perceptions of an individual or specific group. For example, in a lab setting participants rated people with facial anomalies negatively on numerous traits (Workman et al., 2021). In order to combat this negative first impression a number of targeted interventions have been assessed. Researchers found that exposure to positive messages or messages from the target about overcoming adversity, presented by either video or audio, significantly increased positive perception of the target (Stone & Fisher, 2020). This was in contrast to participants in a control condition who were merely exposed to a picture of the target. This work suggests that impressions based off of easily discernible facial features may be able to be consciously updated, becoming more positive with exposure and education regarding a target group.

The effectiveness of educating a participant about the dangers of first impressions and their inaccuracies has also been utilised to combat biases resulting in impressions of trustworthiness from faces (Jaeger et al., 2020c). Jaeger & colleagues found that sentencing decisions remained based on a target's perceived trustworthiness even after participants received explicit warnings about the dangers and inaccuracies of first impressions beforehand. It was found that the level of bias in participant's responses actually increased due to one intervention whereby participants were asked to base their decision first on evidence alone, and then able to change their verdict after seeing a picture of the defendant. This work reflects the difficulty interventions face in attempts at counteracting initial impressions. This may be due to the widespread and deeply ingrained physiognomic beliefs present in some populations (Jaeger et al., 2019), an explicit belief that may take more than a single session to override. Following on from the work of Jaeger and colleagues (2020), research should investigate the effectiveness of prolonged, rather than single exposure, interventions aimed at teaching people about the dangers and ineffectiveness of first impressions.

The enduring nature and consequences of first impressions despite countervailing evidence (Jaeger et al., 2020c) presents a substantial challenge for any interventions. From a social learning perspective, initial first impressions should be able to be revised based on new diagnostic information. To some extent, this is what research has shown, with initial trustworthiness face-based impressions being updated in light of new information, but only when this information is extreme, diagnostic and from a reliable source (Shen et al., 2020; Shen & Ferguson, 2021). This method of impression updating has been explored as one potential intervention, intended to reduce face-based biases. Chua & Freeman (2021), found that having participants engage in "behavioural counter stereotype training" whereby stereotypically untrustworthy faces were paired with mostly positive behavioural descriptions, and trustworthy with negative descriptions, facial stereotype biases were reduced or eliminated. Training of this type attenuated biases in participants' payment and hiring decisions as well as in their automatic face evaluations, measured via an evaluative priming paradigm (Chua & Freeman, 2021). Studies using similar paradigms to reduce racial bias have reported that the effects of counter stereotype training did not last longer than a few days (Forscher et al., 2019). Therefore, future work should attempt to measure how persistent effects of counter stereotype training are across time and continue to test across contexts. Such interventions might be particularly effective in children before stereotypes become deeply entrenched.

Cultural Input

Chapter 4 explores one way through which children may acquire face-trait mappings is through conversations with their parents or caregivers. Parent language use can influence children's knowledge of social categories (Gelman et al., 2004) and essentialist beliefs about social groups (Rhodes et al., 2012; Segall et al., 2015). Previous research has demonstrated how parent-child conversations can facilitate the reduction of racial biases in children (Perry et al., 2020). Based off this previous literature and the work presented in Chapter 4, the possibility of changing the way parents talk about traits and how this relates to children's physiognomic beliefs and first impressions could be explored as an intervention. Furthermore, with some work suggesting that learning preferences shift away from parents around 6-7 years (Lucas et al., 2017), then any parent led interventions may be especially effective for those in their younger years.

The non-verbal behaviour of others is one source of children's knowledge about others (Brey & Pauker, 2019) and can inform their first impressions (see Chapter 3). Work with adults has pointed to the influence of similar types of cultural input on first impressions, with the viewing of racially biased non-verbal cues in the media shown to help shape racial attitudes (Castelli et al., 2012; Weisbuch et al., 2009; Willard et al., 2015). However little work has investigated the possibility of changing this type of cultural input as an intervention to help counter-act harmful first impressions present in a community. As work has mainly focused on representations of gender and race in the media (Dale et al., 2016; Dixon, 2017; Gestos et al., 2018; Weisbuch et al., 2009) a focus for future work could be to identify trends whereby other facial characteristics (e.g. crooked nose) are frequently paired with personality traits and whether this predicts children's first impressions regarding targets who share those features. Once identified targeted interventions could be undertaken whereby children are exposed to characters who possess these features performing positive behaviours. Changes in children's impressions towards novel targets with similar facial features could then be measured.

As well as exploring interventions aiming to both target and change first impressions, it is important to consider interventions designed to prevent the acquisition of inaccurate face-trait mappings. Social media is one potential source of inaccurate face-trait mappings. A large amount of the information shared online is misleading or even inaccurate (Shahi et al., 2020) and fake news appears to spread faster and further online than stories verified to be true (Vosoughi et al., 2018). Sites such as Twitter attempt to ameliorate the many negative consequences of fake news by adding a “disputed tag” to flag inaccurate or distrusted news stories (Allcott et al., 2019; Roth & Pickles, 2020). However, work by (Baum & Rahman, 2021a) has shown that when participants are exposed to headlines from trusted and distrusted websites, the credibility of the site does not modulate social judgements. Instead, participants form first impressions of others’ faces based on the valence of the information associated with the target even when the source of the information is distrusted (Baum & Rahman, 2021a, 2021b). These results highlight social media interventions as one potential key area that future work on first impressions should tackle, with the effectiveness of other measures to counteract inaccurate impressions, like disputed tags, taken in to account. Furthermore, given that first impressions are often generalised to individuals who physically resemble the original targets of learning (FeldmanHall et al., 2018; Lee et al., 2021; Verosky & Todorov, 2013), the negative influence of misinformation and ability to effect judgements of more than just the target should be studied.

Conclusion

Previous work on first impressions has demonstrated their pervasiveness (Jones et al., 2021) and potentially negative consequences (Todorov et al., 2015). On account of this, it is important for work to investigate the mechanisms behind first impressions and to answer questions regarding their development. I hope the research presented in my doctoral thesis represents some of the ways social and developmental psychology can start addressing these questions and can help inform future theories surrounding the origins of first impressions.

References

Adams, R. B. Jr., Garrido, C. O., Albohn, D. N., Hess, U., & Kleck, R. E. (2016). What Facial Appearance Reveals Over Time: When Perceived Expressions in Neutral Faces Reveal Stable Emotion Dispositions. *Frontiers in Psychology*, *0*(JUN), 986. https://doi.org/10.3389/FPSYG.2016.00986

Alaei, R., & Rule, N. O. (2019). People Can Accurately (But Not Adaptively) Judge Strangers' Antigay Prejudice from Faces. *Journal of Nonverbal Behavior*, *43*(3), 397–409. https://doi.org/10.1007/s10919-019-00305-2

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

Albohn, D. N., Brandenburg, J. C., & Adams, R. B. (2019). Perceiving Emotion in the "Neutral" Face: A Powerful Mechanism of Person Perception. In *The Social Nature of Emotion Expression* (pp. 25–47). Springer, Cham. https://doi.org/10.1007/978-3-030-32968-6_3

Albright, L., Kenny, D. A., & Fang, X. (1997). Cross-Cultural Consensus in Personality Judgments Developmental Interpersonal Percepiton View project Analysis of Dyadic and Group Data View project. *Article in Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.72.3.558

Alguacil, S., Tudela, P., & Ruz, M. (2015). Ignoring facial emotion expressions does not eliminate their influence on cooperation decisions. *Psicologica*, *36*(2), 309–335.

Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research and Politics*, *6*(2). https://doi.org/10.1177/2053168019848554

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

Amir, D., & McAuliffe, K. (2020). Cross-cultural, developmental psychology: integrating approaches and key insights. *Evolution and Human Behavior*, *41*(5), 430–444. https://doi.org/10.1016/J.EVOLHUMBEHAV.2020.06.006

Anderson, E., Siegel, E., White, D., & Barrett, L. F. (2012). Out of sight but not out of mind: unseen affective faces influence evaluations and social impressions. *Emotion (Washington, D.C.)*, *12*(6), 1210–1221. https://doi.org/10.1037/a0027514

Antonakis, J. (2011). Predictors of leadership : The usual suspects and the suspect traits. *Sage Handbook of Leadership*, 269–285.

Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play! *Science*, *323*(5918), 1183. https://doi.org/10.1126/science.1167748

Baccolo, E., & Macchi Cassia, V. (2019). Individual differences in perceptual sensitivity and representation of facial signals of trustworthiness. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(2), 224–236. https://doi.org/10.1037/xhp0000601

Baccolo, E., & Macchi Cassia, V. (2020). Age-Related Differences in Sensitivity to Facial Trustworthiness: Perceptual Representation and the Role of Emotional Development. *Child Development*, *91*(5), 1529–1547. https://doi.org/10.1111/cdev.13340

Baccolo, E., Quadrelli, E., & Macchi Cassia, V. (2021). Neural sensitivity to trustworthiness cues from realistic face images is associated with temperament: an electrophysiological study with 6-month-old infants. *Social Neuroscience*. https://doi.org/10.1080/17470919.2021.1976271

Baum, J., & Rahman, R. A. (2021a). Emotional news affects social judgments independent of perceived media credibility. *Social Cognitive and Affective Neuroscience*, *16*(3), 280–291. https://doi.org/10.1093/scan/nsaa164

Baum, J., & Rahman, R. (2021b). Negative news dominates fast and slow brain responses and social judgments even after source credibility evaluation. *NeuroImage*, *244*, 118572. https://doi.org/10.1016/J.NEUROIMAGE.2021.118572

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

Bargh, J. A. (1994). The four horsemen of automaticity: Intention, awareness, efficiency, and control as separate issues. In *Handbook of social cognition* (pp. 1–40). https://philpapers.org/rec/BARTFH-3

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

Bazzini, D., Curtin, L., Joslin, S., Regan, S., & Martz, D. (2010). Do Animated Disney Characters Portray and Promote the Beauty-Goodness Stereotype? *Journal of Applied Social Psychology*, *40*(10), 2687–2709. https://doi.org/10.1111/j.1559-1816.2010.00676.x

Behling, D. U., & Williams, E. A. (2016). Influence of Dress on Perception of Intelligence and Expectations of Scholastic Achievement: *Http://Dx.Doi.Org/10.1177/0887302X9100900401*, *9*(4), 1–7. https://doi.org/10.1177/0887302X9100900401

Berry, D. S., & McArthur, L. Z. (1986). Perceiving Character in Faces. The Impact of Age-Related Craniofacial Changes on Social Perception. In *Psychological Bulletin* (Vol. 100, Issue 1, pp. 3–18). https://doi.org/10.1037/0033-2909.100.1.3

Berry, D. S., & Zebrowitz-McArthur, L. (1988). What's in a Face? *Personality and Social Psychology Bulletin*, *14*(1), 23–33. https://doi.org/10.1177/0146167288141003

Bjornsdottir, R. T., & Rule, N. O. (2017). The visibility of social class from facial cues. *Journal of Personality and Social Psychology*, *113*(4), 530–546. https://doi.org/10.1037/pspa0000091

Blaskovits, B., & Bennell, C. (2019). Are we Revealing Hidden Aspects of our Personality When we Walk? *Journal of Nonverbal Behavior*, *43*(3), 329–356. https://doi.org/10.1007/S10919-019-00302-5

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2015). Face-ism and kernels of truth in facial inferences. In *Trends in Cognitive Sciences* (Vol. 19, Issue 8, pp. 421–422). https://doi.org/10.1016/j.tics.2015.05.002

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2017). Can We Detect Cooperators by Looking at Their Face? *Current Directions in Psychological Science*, *26*(3), 276–281. https://doi.org/10.1177/0963721417693352

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Brey, E., & Pauker, K. (2019). Teachers' nonverbal behaviors influence children's stereotypic beliefs. *Journal of Experimental Child Psychology*, *188*, 104671. https://doi.org/10.1016/j.jecp.2019.104671

Brey, E., & Shutts, K. (2015). Children Use Nonverbal Cues to Make Inferences About Social Power. *Child Development*, *86*(1), 276–286. https://doi.org/10.1111/cdev.12334

Brosseau-Liard, P. E., & Birch, S. A. J. (2010). "I bet you know more and are nicer too!": What children infer from others' accuracy. *Developmental Science*. https://doi.org/10.1111/j.1467-7687.2009.00932.x

Brown, R. D., & Bassili, J. N. (2002). Spontaneous trait associations and the case of the superstitious banana. *Journal of Experimental Social Psychology*, *38*(1), 87–92. https://doi.org/10.1006/jesp.2001.1486

Brusattin, L. (2012). Candidate Visual Appearance as a Shortcut for Both Sophisticated and Unsophisticated Voters: Evidence from a Spanish Online Study. *International Journal of Public Opinion Research*, *24*(1), 1–20. https://doi.org/10.1093/ijpor/edr040

Bülthoff, I., & Zhao, M. (2021). Average faces: How does the averaging process change faces physically and perceptually?. *Cognition*, *216*, 104867.

Canace, T. G., Cianci, A. M., Liu, X. K., & Tsakumis, G. T. (2020). Paid for looks when others are looking: CEO facial traits, compensation, and corporate visibility. *Journal of Business Research*, *115*, 85-100.

Carlston, D. E., & Skowronski, J. J. (1994). Savings in the Relearning of Trait Information as Evidence for Spontaneous Inference Generation. *Journal of Personality and Social Psychology*, *66*(5), 840–856. https://doi.org/10.1037/0022-3514.66.5.840

Carlston, D. E., Skowronski, J. J., & Sparks, C. (1995). Savings in Relearning: II. On the Formation of Behavior-Based Trait Associations and Inferences. *Journal of*

Personality and Social Psychology, *69*(3), 420–436. https://doi.org/10.1037/0022-3514.69.3.429

Carney, D. R., Colvin, C. R., & Hall, J. A. (2007). A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality*. https://doi.org/10.1016/j.jrp.2007.01.004

Castelli, L., Carraro, L., Pavan, G., Murelli, E., & Carraro, A. (2012). The Power of the Unsaid: The Influence of Nonverbal Cues on Implicit Attitudes. *Journal of Applied Social Psychology*, *42*(6), 1376–1393. https://doi.org/10.1111/j.1559-1816.2012.00903.x

Caulfield, F., Ewing, L., Bank, S., & Rhodes, G. (2016). Judging trustworthiness from faces: Emotion cues modulate trustworthiness judgments in young children. *British Journal of Psychology*, *107*(3), 503–518. https://doi.org/10.1111/bjop.12156

Chalik, L., & Rhodes, M. (2015). The Communication of Naïve Theories of the Social World in Parent–Child Conversation. *Journal of Cognition and Development*, *16*(5), 719–741. https://doi.org/10.1080/15248372.2014.949722

Chang, L. J., Doll, B. B., van 't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, *61*(2), 87–105. https://doi.org/10.1016/J.COGPSYCH.2010.03.001

Charlesworth, T. E. S., & Banaji, M. R. (2019). Face-Trait and Face-Race Cues in Adults' and Children's Social Evaluations. *Social Cognition*, *37*(4), 357–388. https://doi.org/10.1521/soco.2019.37.4.357

Charlesworth, T. E., Hudson, S. K. T., Cogsdill, E. J., Spelke, E. S., & Banaji, M. R. (2019). Children use targets' facial appearance to guide and predict social behavior. *Developmental psychology*, *55*(7), 1400.

Chronaki, G., Hadwin, J. A., Garner, M., Maurage, P., & Sonuga-Barke, E. J. S. (2015). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. *British Journal of Developmental Psychology*, *33*(2), 218–236. https://doi.org/10.1111/bjdp.12075

Chua, K.-W., & Freeman, J. B. (2021). Facial Stereotype Bias Is Mitigated by Training. *Social Psychological and Personality Science*, *12*(7), 1335–1344. https://doi.org/10.1177/1948550620972550

Cikara, M., & van Bavel, J. J. (2014). The neuroscience of intergroup relations: an integrative review. *Perspectives on Psychological Science*, *9*(3), 245–274. https://doi.org/10.1177/1745691614527464

Cogsdill, E. J., & Banaji, M. R. (2015). Face-trait inferences show robust child-adult agreement: Evidence from three types of faces. *Journal of Experimental Social Psychology*, *60*, 150–156. https://doi.org/10.1016/j.jesp.2015.05.007

Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring Character From Faces: A Developmental Study. *Psychological Science*, *25*(5), 1132–1139. https://doi.org/10.1177/0956797614523297

Collins, A. F. (1999). The Enduring Appeal of Physiognomy: Physical Appearance as a Sign of Temperament, Character, and Intelligence. *History of Psychology*. https://doi.org/10.1037/1093-4510.2.4.251

Collova, J. R., Jeffery, L., Rhodes, G., Bothe, E., & Sutherland, C. A. (2020a). Evidence for a kernel of truth in children's facial impressions of children's niceness but not shyness. *Developmental Psychology*, 283.

Collova, J. R., Sutherland, C. A., Jeffery, L., Bothe, E., & Rhodes, G. (2020b). Adults' facial impressions of children's niceness, but not shyness, show modest accuracy. *Quarterly Journal of Experimental Psychology*, 174702182095757. https://doi.org/10.1177/1747021820957575

Collova, J. R., Sutherland, C. A. M., & Rhodes, G. (2019). Testing the functional basis of first impressions: Dimensions for children's faces are not the same as for adults' faces. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspa0000167

Cook, R., & Over, H. (2020). A learning model can explain both shared and idiosyncratic first impressions from faces. In *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 117, Issue 28, pp. 16112–16113). National Academy of Sciences. https://doi.org/10.1073/pnas.2008816117

Croley, J. A., Reese, V., & Wagner, R. F. (2017). Dermatologic features of classic movie villains: The face of evil. *JAMA Dermatology*, *153*(6), 559–564. https://doi.org/10.1001/jamadermatol.2016.5979

Dale, L. P., Higgins, B. E., Pinkerton, N., Couto, M., Mansolillo, V., Weisinger, N., & Flores, M. (2016). Princess Picture Books: Content and Messages. *Journal of Research in Childhood Education*, *30*(2), 185–199. https://doi.org/10.1080/02568543.2016.1143415

Dalrymple, K. A., Gomez, J., & Duchaine, B. (2013). The Dartmouth Database of Children's Faces: Acquisition and Validation of a New Face Stimulus Set. *PLoS ONE*, *8*(11), e79131. https://doi.org/10.1371/journal.pone.0079131

Davis, J. T. M., Robertson, E., Lew-Levy, S., Neldner, K., Kapitany, R., Nielsen, M., & Hines, M. (2021). Cultural Components of Sex Differences in Color Preference. *Child Development*, *92*(4), 1574–1589. https://doi.org/10.1111/CDEV.13528

DeBruine, L. (2017). Webmorph (Version v0. 0.0. 9001). *Webmorph (Version v0.0.0.9001)*, *Zenodo*.

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

De Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, *14*(3), 289–311. https://doi.org/10.1080/026999300378824

DeJesus, J. M., Shutts, K., & Kinzler, K. D. (2018). Mere social knowledge impacts children's consumption and categorization of foods. *Developmental Science*, *21*(5), e12627. https://doi.org/10.1111/DESC.12627

de Neys, W., Hopfensitz, A., & Bonnefon, J.-F. (2015). Adolescents gradually improve at detecting trustworthiness from the facial features of unknown adults. https://doi.org/10.1016/j.joep.2015.01.002

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5.

Diekman, A. B., & Murnen, S. K. (2004). Learning to Be Little Women and Little Men: The Inequitable Gender Equality of Nonsexist Children's Literature. *Sex Roles*, *50*(5–6), 373–385. https://doi.org/10.1023/b:sers.0000018892.26527.ea

Dilger, A., Muller, J., & Muller, M. (2017). Is Trustworthiness Written on the Face? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2930064

Dill, K. E., & Thill, K. P. (2007). Video game characters and the socialization of gender roles: Young people's perceptions mirror sexist media depictions. *Sex Roles*, *57*(11–12), 851–864. https://doi.org/10.1007/s11199-007-9278-1

Dixon, T. L. (2017). Good Guys Are Still Always in White? Positive Change and Continued Misrepresentation of Race and Crime on Local Television News. *Communication Research*. https://doi.org/10.1177/0093650215579223

Dobrow, J. R., & Gidney, C. L. (1998). The Good, the Bad, and the Foreign: The Use of Dialect in Children's Animated Television: *The Annals of the American Academy of Politicl and Social Science*, *557*(1), 105–119. https://doi.org/10.1177/0002716298557000009

Dunham, Y., Baron, A. S., & Banaji, M. R. (2016). The development of implicit gender attitudes. *Developmental Science*, *19*(5), 781–789. https://doi.org/10.1111/desc.12321

Dunham, Y., Stepanova, E. v., Dotsch, R., & Todorov, A. (2015). The development of race-based perceptual categorization: skin color dominates early category judgments. *Developmental Science*, *18*(3), 469–483. https://doi.org/10.1111/desc.12228

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

Eggleston, A., Flavell, J. C., Tipper, S. P., Cook, R., & Over, H. (2021). Culturally learned first impressions occur rapidly and automatically and emerge early in development. *Developmental Science*, *24*(2). https://doi.org/10.1111/desc.13021

Engel-Yeger, B., & Jarus, T. (2008). Cultural and gender effects on Israeli children's preferences for activities. *Canadian Journal of Occupational Therapy*, *75*(3), 139–148. https://doi.org/10.1177/000841740807500305

Erdfelder, E., Faul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Evans, E. C. (1969). Physiognomics in the Ancient World. *Transactions of the American Philosophical Society*, *59*(5), 1. https://doi.org/10.2307/1006011

Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2015). Perceived trustworthiness of faces drives trust behaviour in children. *Developmental Science*, *18*(2), 327–334.

Ewing, L., Sutherland, C. A. M., & Willis, M. L. (2019). Children show adult-like facial appearance biases when trusting others. *Developmental Psychology*. https://doi.org/10.1037/dev0000747

Fein, G. G. (1975). Children's sensitivity to social contexts at 18 months of age. *Developmental Psychology*, *11*(6), 853–854. https://doi.org/10.1037/0012-1649.11.6.853

FeldmanHall, O., Dunsmoor, J. E., Tompary, A., Hunter, L. E., Todorov, A., & Phelps, E. A. (2018). *Stimulus generalization as a mechanism for learning to trust*. *115*(7), E1690–E1697. https://doi.org/10.1073/pnas.1715227115

Filkuková, P., & Jørgensen, M. (2020). How to pose for a professional photo: The effect of three facial expressions on perception of competence of a software developer. *Australian Journal of Psychology*, *72*(3), 257-266.

Fleischmann, A., Lammers, J., Stoker, J. I., & Garretsen, H. (2019). You Can Leave Your Glasses on: Glasses Can Increase Electoral Success. *Social Psychology*, *50*(1), 38–52. https://doi.org/10.1027/1864-9335/a000359

Flowe, H. D. (2012). Do Characteristics of Faces That Convey Trustworthiness and Dominance Underlie Perceptions of Criminality? *PLoS ONE*, *7*(6), e37253. https://doi.org/10.1371/journal.pone.0037253

Fong, F. T. K., Imuta, K., Redshaw, J., & Nielsen, M. (2021). The digital social partner: Preschool children display stronger imitative tendency in screen-based than live learning. *Human Behavior and Emerging Technologies*, hbe2.280. https://doi.org/10.1002/HBE2.280

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A Meta-Analysis of Procedures to Change Implicit Measures. *Journal of Personality and Social Psychology*, *117*(3), 522. https://doi.org/10.1037/PSPA0000160

Funk, F., & Todorov, A. (2013). *Criminal Stereotypes in the Courtroom: Facial Tattoos Affect Guilt and Punishment Differently*. https://doi.org/10.1037/a0034736

Funk, F., Walker, M., & Todorov, A. (2016). Modelling perceptions of criminality and remorse from faces using a data-driven computational approach. *Cognition and Emotion*, 1–13. https://doi.org/10.1080/02699931.2016.1227305

Galton, F. (2007). Inquiries into human faculty and its development, 1883. In *Readings in the history of psychology.* https://doi.org/10.1037/11304-032

Gao, X., & Maurer, D. (2009). Influence of intensity on children's sensitivity to happy, sad, and fearful facial expressions. *Journal of Experimental Child Psychology*, *102*(4), 503–521. https://doi.org/10.1016/j.jecp.2008.11.002

Gelman, S. A., Taylor, M. G., Nguyen, S. P., & Bigler, R. S. (2004). Mother-Child Conversations about Gender: Understanding the Acquisition of Essentialist Beliefs. *Mother-Child Conversations about Gender: Understanding the Acquisition of Essentialist Beliefs*, *69*(1), 1–142.

Gendron, M., Roberson, D., Vyver, J. M. van der, & Barrett, L. F. (2014). Perceptions of Emotion from Facial Expressions are Not Culturally Universal: Evidence from a Remote Culture. *Emotion (Washington, D.C.)*, *14*(2), 251. https://doi.org/10.1037/A0036052

Germine, L., Russell, R., Bronstad, P. M., Blokland, G. A. M., Smoller, J. W., Kwok, H., Anthony, S. E., Nakayama, K., Rhodes, G., & Wilmer, J. B. (2015). Individual Aesthetic Preferences for Faces Are Shaped Mostly by Environments, Not Genes. *Current Biology*. https://doi.org/10.1016/j.cub.2015.08.048

Gestos, M., Smith-Merry, J., & Campbell, A. (2018). Representation of Women in Video Games: A Systematic Review of Literature in Consideration of Adult Female Wellbeing. In *Cyberpsychology, Behavior, and Social Networking* (Vol. 21, Issue 9, pp. 535–541). Cyberpsychol Behav Soc Netw. https://doi.org/10.1089/cyber.2017.0376

Graham, J. R., Harvey, C. R., & Puri, M. (2017). A Corporate Beauty Contest. *Management Science*, *63*(9), 3044–3056. https://doi.org/10.1287/mnsc.2016.2484

Gray, R. T. (2004). *About Face: German Physiognomic Thought from Lavater to Auschwitz*. Wayne State University Press. https://doi.org/10.2307/20466881

Guajardo, N. R., Snyder, G., & Petersen, R. (2009). Relationships among Parenting Practices, Parental Stress, Child Behaviour, and Children's Social-Cognitive Development. *Infant and Child Development*, *18*(1), 37–60. https://doi.org/10.1002/icd.578

Gutiérrez-García, A., Beltrán, D., & Calvo, M. G. (2018). Facial attractiveness impressions precede trustworthiness inferences: lower detection thresholds and faster decision latencies. *Https://Doi.Org/10.1080/02699931.2018.1444583*, *33*(2), 378–385. https://doi.org/10.1080/02699931.2018.1444583

Hackel, L. M., Doll, B. B., & Amodio, D. M. (2015). Instrumental learning of traits versus rewards: Dissociable neural correlates and effects on choice. *Nature Neuroscience*, *18*(9), 1233–1235. https://doi.org/10.1038/nn.4080

Han, C., Wang, H., Hahn, A. C., Fisher, C. I., Kandrik, M., Fasolt, V., Morrison, D. K., Lee, A. J., Holzleitner, I. J., DeBruine, L. M., & Jones, B. C. (2018). Cultural differences in preferences for facial coloration. *Evolution and Human Behavior*, *39*(2), 154–159. https://doi.org/10.1016/j.evolhumbehav.2017.11.005

Hansen, J. (2009). The Art and Science of Reading Faces: Strategies of Racist Cinema in the Third Reich. *Shofar An Interdisciplinary Journal of Jewish Studies*. https://doi.org/10.1353/sho.0.0528

Haselhuhn, M. P., Wong, E. M., & Ormiston, M. E. (2013). Self-Fulfilling Prophecies as a Link between Men's Facial Width-to-Height Ratio and Behavior. *PLOS ONE*, *8*(8), e72259. https://doi.org/10.1371/JOURNAL.PONE.0072259

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

Hehman, E., Graber, E. C., Hoffman, L. H., & Gaertner, S. L. (2012). Warmth and competence: A content analysis of photographs depicting American presidents. *Psychology of Popular Media Culture*, *1*(1), 46–52. https://doi.org/10.1037/a0026513

Hehman, E., Sutherland, C. A., Flake, J. K., & Slepian, M. L. (2017). The unique contributions of perceiver and target characteristics in person perception. *Journal of personality and social psychology*, *113*(4), 513.

Hellström, Å., & Tekle, J. (1994). Person perception through facial photographs: Effects of glasses, hair, and beard on judgments of occupation and personal qualities. *European Journal of Social Psychology*, *24*(6), 693–705.

Hennefield, L., & Markson, L. (2017). Four-year-old Children Align their Preferences with those of their Peers. *Collabra: Psychology*, *3*(1). https://doi.org/10.1525/COLLABRA.89

Henrich, J., & Broesch, J. (2011). On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1139-1148.

Heyes, C. (2018). Cognitive gadgets: The cultural evolution of thinking. In *Harvard University Press*. Harvard University Press. https://doi.org/10.4159/9780674985155

Heyes, C., & Pearce, J. M. (2015). Not-so-social learning strategies. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1802), 20141709.

Hollis-Sawyer, L., & Cuevas, L. (2013). Mirror, Mirror on the Wall: Ageist and Sexist Double Jeopardy Portrayals in Children's Picture Books. *Educational Gerontology*, *39*(12), 902–914. https://doi.org/10.1080/03601277.2013.767650

Hong, S., Suk, H. W., Choi, Y., & Na, J. (2021). Face-Based Judgments: Accuracy, Validity, and a Potential Underlying Mechanism. *Psychological Science*. https://doi.org/10.1177/09567976211000308

Ilardi, V. (2007). Renaissance Vision from Spectacles to Telescopes. *American Philosophical Society*, *259*.

Jaeger, B., Evans, A., Stel, M., & van Beest, I. (2020b). *Who judges a book by its cover? Belief in physiognomy and reliance on trustworthiness impressions from faces*. https://doi.org/10.31234/osf.io/8dq4x

Jaeger, B., Sleegers, W. W. A., Stern, J., Penke, L., & Jones, A. L. (2020a). *The accuracy and meta-accuracy of personality impressions from faces*. https://osf.io/tus85/download

Jaeger, B., Todorov, A. T., Evans, A. M., & van Beest, I. (2020c). Can we reduce facial biases? Persistent effects of facial trustworthiness on sentencing decisions. *Journal of Experimental Social Psychology*, *90*, 104004. https://doi.org/10.1016/j.jesp.2020.104004

Jessen, S., & Grossmann, T. (2016). Neural and Behavioral Evidence for Infants' Sensitivity to the Trustworthiness of Faces. *Journal of Cognitive Neuroscience*, *28*(11), 1728–1736. https://doi.org/10.1162/jocn_a_00999

Jessen, S., & Grossmann, T. (2017). Neural evidence for the subliminal processing of facial trustworthiness in infancy. *Neuropsychologia*, 1–8. https://doi.org/10.1016/j.neuropsychologia.2017.04.025

Jessen, S., & Grossmann, T. (2019). Neural evidence for the subliminal processing of facial trustworthiness in infancy. *Neuropsychologia*, *126*, 46–53. https://doi.org/10.1016/j.neuropsychologia.2017.04.025

Johnson, M. H. (2005). Subcortical face processing. In *Nature Reviews Neuroscience* (Vol. 6, Issue 10, pp. 766–774). Nature Publishing Group. https://doi.org/10.1038/nrn1766

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

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

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

Karnadewi, F., & Lipp, O. (2011). The Processing of Invariant and Variant Face Cues in the Garner Paradigm. *Emotion*, *11*(3), 563–571. https://doi.org/10.1037/a0021333

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(30), 12577–12580. https://doi.org/10.1073/pnas.0705345104

Klatt, T., Phelps, M., Maltby, J., Smailes, H. L., Ryder, H. L., Humphries, J. E., & Flowe, H. D. (2016). *Looking bad: Inferring criminality after 100 milliseconds*.

Kleiner, M., Brainard, D. H., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(ECVP Abstract Supplement).

Klein, H., & Shiffman, K. S. (2006). Messages about physical attractiveness in animated cartoons. *Body Image*, *3*(4), 353–363. https://doi.org/10.1016/j.bodyim.2006.08.001

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1738), 2698–2704. https://doi.org/10.1098/RSPB.2012.0311

Knight, S., Lavan, N., Torre, I., & McGettigan, C. (2021). The influence of perceived vocal traits on trusting behaviours in an economic game. *Quarterly Journal of Experimental Psychology*, *74*(10), 1747–1754. https://doi.org/10.1177/17470218211010144

Kocsor, F., & Bereczkei, T. (2017). Evaluative conditioning leads to differences in the social evaluation of prototypical faces. *Personality and Individual Differences*, *104*, 215–219. https://doi.org/10.1016/J.PAID.2016.08.007

Kocsor, F., Kozma, L., Neria, A. L., Jones, D. N., & Bereczkei, T. (2019). Arbitrary signals of trustworthiness – social judgments may rely on facial expressions even with experimentally manipulated valence. *Heliyon*, *5*(5), e01736. https://doi.org/10.1016/J.HELIYON.2019.E01736

Krys, K., Hansen, K., Xing, C., Szarota, P., & Yang, M. miao. (2014). Do Only Fools Smile at Strangers? Cultural Differences in Social Perception of Intelligence of Smiling Individuals. *Journal of Cross-Cultural Psychology*, *45*(2), 314–321. https://doi.org/10.1177/0022022113513922

Krys, K., -Melanie Vauclair, C., Capaldi, C. A., Lun, V. M. C., Bond, M. H., Domínguez-Espinosa, A., Torres, C., Lipp, O. v., Manickam, L. S. S., Xing, C., Antalíková, R., Pavlopoulos, V., Teyssier, J., Hur, T., Hansen, K., Szarota, P., Ahmed, R. A., Burtceva, E., Chkhaidze, A., … Yu, A. A. (2016). Be Careful Where You Smile: Culture Shapes Judgments of Intelligence and Honesty of Smiling Individuals. *Journal of Nonverbal Behavior*, *40*(2), 101–116. https://doi.org/10.1007/s10919-015-0226-4

LaBounty, J., Wellman, H. M., Olson, S., Lagattuta, K., & Liu, D. (2008). Mothers' and fathers' use of internal state talk with their young children. *Social Development*, *17*(4), 757–775. https://doi.org/10.1111/j.1467-9507.2007.00450.x

Lane, J. D., Wellman, H. M., & Gelman, S. A. (2013). Informants' traits weigh heavily in young children's trust in testimony and in their epistemic inferences. *Child Development*. https://doi.org/10.1111/cdev.12029

Langlois, J. H., & Roggman, L. A. (1990). Attractive Faces Are Only Average. *Psychological Science*, *1*(2), 115–121. https://doi.org/10.1111/j.1467-9280.1990.tb00079.x

LaTosky, S. (2006). Reflections on the lip-plates of Mursi women as a source of stigma and self-esteem. *Perils of Face: Essays on Cultural Contact, Respect and Self-Esteem in Southern Ethiopia*.

Lavan, N., Mileva, M., Burton, A. M., Young, A. W., & McGettigan, C. (2021). Trait evaluations of faces and voices: Comparing within- and between-person variability. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/XGE0001019

Lavan, N., Mileva, M., & Mcgettigan, C. (2020). How does familiarity with a voice affect trait judgements? *British Journal of Psychology*. https://doi.org/10.1111/bjop.12454

Lawson, C., Lenz, G. S., Baker, A., & Myers, M. (2010). Looking Like a Winner: Candidate Appearance and Electoral Success in New Democracies Looking Like a winner Candidate appearance and electoral Success in new Democracies. *World Politics*, *62*(4), 561–593. https://doi.org/10.1017/S0043887110000195

Lee, R., Flavell, J. C., Tipper, S. P., Cook, R., & Over, H. (2021). Spontaneous first impressions emerge from brief training. *Scientific Reports*.

Lenz, G. S., & Lawson, C. (2011). Looking the part: Television leads less informed citizens to vote based on candidates' appearance. *American Journal of Political Science*, *55*(3), 574–589. https://doi.org/10.1111/j.1540-5907.2011.00511.x

Lewis, M., Borkenhagen, M. C., Converse, E., Lupyan, G., & Seidenberg, M. S. (2020). *What might books be teaching young children about gender?* https://doi.org/10.31234/OSF.IO/NTGFE

Lin, C., Adolphs, R., & Alvarez, R. M. (2018). Inferring Whether Officials Are Corruptible From Looking at Their Faces. *Psychological Science*, *29*(11), 1807–1823. https://doi.org/10.1177/0956797618788882

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

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

Li, Q., Heyman, G. D., Mei, J., & Lee, K. (2017). *Judging a Book by Its Cover: Children's Facial Trustworthiness as Judged by Strangers Predicts Their Real-World Trustworthiness and Peer Relationships*. https://doi.org/10.1111/cdev.12907

Little, A. C. (2014). Facial appearance and leader choice in different contexts: Evidence for task contingent selection based on implicit and learned face-behaviour/face-ability associations. *The Leadership Quarterly*, *25*(5), 865–874. https://doi.org/10.1016/J.LEAQUA.2014.04.002

Little, A. C., & Hancock, P. J. B. (2002). The role of masculinity and distinctiveness in judgments of human male facial attractiveness. *British Journal of Psychology*, *93*(4), 451–464. https://doi.org/10.1348/000712602761381349

Little, A. C., Jones, B. C., DeBruine, L. M., & Dunbar, R. I. M. (2013). Accuracy in discrimination of self-reported cooperators using static facial information. *Personality and Individual Differences*. https://doi.org/10.1016/j.paid.2012.10.018

Li, Y., Jiao, X., Liu, Y., Tse, C., & Dong, Y. (2020). Age differences in facial trustworthiness judgement based on multiple facial cues. *British Journal of Psychology*, bjop.12472. https://doi.org/10.1111/bjop.12472

Lippi-Green, Rosina. (2012). *English with an accent : language, ideology and discrimination in the United States*. 101–129.

Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of Young Children's Trait Understanding: Behavior-to-Trait Inferences and Trait-to-Behavior Predictions. *Child Development*, *78*(5), 1543–1558.

Lucas, A. J., Burdett, E. R. R., Burgess, V., Wood, L. A., McGuigan, N., Harris, P. L., & Whiten, A. (2017). The Development of Selective Copying: Children's Learning From an Expert Versus Their Mother. *Child Development*, *88*(6), 2026–2042. https://doi.org/10.1111/CDEV.12711

Lundqvist, D., Flykt, A., & Ohman, A. (1998). The Karolinska directed emotional faces (KDEF). In *CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet*. https://doi.org/10.1017/S0048577299971664

Ma, D. S., Correll, J., & Wittenbrink, B. (2015a). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Ma, F., Xu, F., & Luo, X. (2015b). Children's and adults' judgments of facial trustworthiness: The relationship to facial attractiveness. *Perceptual and Motor Skills*, *121*(1), 179–198. https://doi.org/10.2466/27.22.PMS.121c10x1

Ma, F., Xu, F., & Luo, X. (2016). Children's Facial Trustworthiness Judgments: Agreement and Relationship with Facial Attractiveness. *Frontiers in Psychology*, *7*, 499. https://doi.org/10.3389/fpsyg.2016.00499

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

Maran, T., Liegl, S., Moder, S., Kraus, S., & Furtner, M. (2021). Clothes make the leader! How leaders can use attire to impact followers' perceptions of charisma and approval. *Journal of Business Research*, *124*, 86–99. https://doi.org/10.1016/J.JBUSRES.2020.11.026

Martin, D. S. (1978). Person perception and real-life electoral behaviour. *Australian Journal of Psychology*, *30*(3), 255–262. https://doi.org/10.1080/00049537808256378

Masur, E. F., & Gleason, J. B. (1980). Parent–child interaction and the acquisition of lexical information during play. *Developmental Psychology*, *16*(5), 404.

Mazur, A., Mazur, J., & Keating, C. (1984). Military Rank Attainment of a West Point Class: Effects of Cadets' Physical Features. *American Journal of Sociology*, *90*(1), 125–150. https://doi.org/10.1086/228050

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLOS ONE*, *9*(3), e90779. https://doi.org/10.1371/JOURNAL.PONE.0090779

McDade-Montez, E., Wallander, J., & Cameron, L. (2017). Sexualization in U.S. Latina and White Girls' Preferred Children's Television Programs. *Sex Roles*, *77*(1–2). https://doi.org/10.1007/s11199-016-0692-0

Menegatti, M., Pireddu, S., Crocetti, E., Moscatelli, S., & Rubini, M. (2021). The Ginevra de' Benci Effect: Competence, Morality, and Attractiveness Inferred From Faces Predict Hiring Decisions for Women. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/FPSYG.2021.658424

Mileva, M., Tompkinson, J., Watt, D., & Burton, A. M. (2017). Audiovisual Integration in Social Evaluation. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/xhp0000439

Miller, M. K., & Summers, A. (2007). Gender differences in video game characters' roles, appearances, and attire as portrayed in video game magazines. *Sex Roles*, *57*(9–10), 733–742. https://doi.org/10.1007/S11199-007-9307-0

Mondloch, C. J., Gerada, A., Proietti, V., & Nelson, N. L. (2019). The influence of subtle facial expressions on children's first impressions of trustworthiness and dominance is not adult-like. *Journal of Experimental Child Psychology*, *180*, 19–38. https://doi.org/10.1016/J.JECP.2018.12.002

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

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*(4), 495–500. https://doi.org/10.1016/0163-6383(93)80007-U

Morris, T. L., Gorham, J., Cohen, S. H., & Huffman, D. (1996). Fashion in the classroom: Effects of attire on student perceptions of instructors in college classes. *Communication Education*, *45*(2), 135–148. https://doi.org/10.1080/03634529609379043

Mueller, U., & Mazur, A. (1996). Facial Dominance of West Point Cadets as a Predictor of Later Military Rank. *Social Forces*, *74*(3), 823. https://doi.org/10.2307/2580383

Nelson, N. L., & Mondloch, C. J. (2017). Adults' and children's perception of facial expressions is influenced by body postures even for dynamic stimuli. *Https://Doi.Org/10.1080/13506285.2017.1301615*, *25*(4–6), 563–574. https://doi.org/10.1080/13506285.2017.1301615

Nesse, R. M. (2006). The Smoke Detector Principle. *Annals of the New York Academy of Sciences*, *935*(1), 75–85. https://doi.org/10.1111/j.1749-6632.2001.tb03472.x

Nielsen, M. (2013). Young Children's Imitative and Innovative Behaviour on the Floating Object Task. *Infant and Child Development*, *22*(1), 44–52. https://doi.org/10.1002/icd.1765

Nielsen, M., Fong, F. T. K., & Whiten, A. (2021). Social learning from media: The need for a culturally diachronic developmental psychology. In *Advances in Child Development and Behavior* (Vol. 61, pp. 317–334). Academic Press Inc. https://doi.org/10.1016/bs.acdb.2021.04.001

Nielsen, M., & Haun, D. (2016). Why developmental psychology is incomplete without comparative and cross-cultural perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686). https://doi.org/10.1098/rstb.2015.0071

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. https://doi.org/10.1016/J.JECP.2017.04.017

Nielsen, M., Mushin, I., Tomaselli, K., & Whiten, A. (2014). Where Culture Takes Hold: "Overimitation" and Its Flexible Deployment in Western, Aboriginal, and Bushmen Children. *Child Development*, *85*(6), 2169–2184. https://doi.org/10.1111/CDEV.12265

Northup, T., & Liebler, C. (2010). The Good, the Bad, and the Beautiful: Beauty ideals on the Disney and Nickelodeon channels. *Journal of Children and Media*, *4*(3), 265–282. https://doi.org/10.1080/17482798.2010.496917

Oh, D. W., Dotsch, R., Porter, J., & Todorov, A. (2019). Gender Biases in Impressions From Faces: Empirical Studies and Computational Models. *Journal of Experimental Psychology: General*, *149*(2). https://doi.org/10.1037/xge0000638

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

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

Olivola, C. Y., Sussman, A. B., Tsetsos, K., Kang, O. E., & Todorov, A. (2012). Republicans prefer Republican-looking leaders: Political facial stereotypes predict candidate electoral success among right-leaning voters. *Social Psychological and Personality Science*, *3*(5), 605–613. https://doi.org/10.1177/1948550611432770

Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-Based Trait Inferences and Voting. *Journal of Nonverbal Behavior*, *34*(2), 83–110. https://doi.org/10.1007/s10919-009-0082-1

Olszanowski, M., Parzuchowski, M., & Szymków, A. (2019). When the smile is not enough: The interactive role of smiling and facial characteristics in forming judgments about trustworthiness and dominance. *Roczniki Psychologiczne*, *22*(1), 35–52. https://doi.org/10.18290/RPSYCH.2019.22.1-3

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

Over, H., & Carpenter, M. (2009). Priming third-party ostracism increases affiliative imitation in children. *Developmental Science*, *12*, 1–8. https://doi.org/10.1111/j.1467-7687.2008.00820.x

Over, H., & Carpenter, M. (2015). Children infer affiliative and status relations from watching others imitate. *Developmental Science*, *18*(6), 917–925. https://doi.org/10.1111/desc.12275

Over, H., & Cook, R. (2018). Where do spontaneous first impressions of faces come from?. *Cognition*, *170*, 190-200.

Over, H., Eggleston, A., & Cook, R. (2020). Ritual and the origins of first impressions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1805), 20190435. https://doi.org/10.1098/rstb.2019.0435

Palmquist, C. M., Cheries, E. W., & DeAngelis, E. R. (2019). Looking smart: Preschoolers' judgements about knowledge based on facial appearance. *British Journal of Developmental Psychology*. https://doi.org/10.1111/bjdp.12303

Palmquist, C. M., & DeAngelis, E. R. (2020). Valence or traits? Developmental change in children's use of facial features to make inferences about others. *Cognitive Development*, *56*, 100948. https://doi.org/10.1016/j.cogdev.2020.100948

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. In *Spatial Vision* (Vol. 10, Issue 4, pp. 437–442). https://doi.org/10.1163/156856897X00366

Peng, Y. (2018). Same candidates, different faces: Uncovering media bias in visual portrayals of presidential candidates with computer vision. *Journal of Communication*, *68*(5), 920-941.

Perry, S., Skinner-Dorkenoo, A., Abaied, J., & Osnaya, A. (2020). Initial Evidence that Parent-Child Conversations About Race Reduce Racial Biases Among White US Children. *PsyArXiv*. https://psyarxiv.com/3xdg8/

Pirchio, S., Passiatore, Y., Panno, A., Maricchiolo, F., & Carrus, G. (2018). A chip off the old block: Parents' subtle ethnic prejudice predicts children's implicit prejudice. *Frontiers in Psychology*, *9*, 1–9. https://doi.org/10.3389/fpsyg.2018.00110

Plachouri, K. M., & Georgiou, S. (2019). Not only a Hollywood trend: the dermatological features of villains in classic and contemporary literature. In *British Journal of Dermatology* (Vol. 181, Issue 3, p. 592). Blackwell Publishing Ltd. https://doi.org/10.1111/bjd.17782

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

Porter, S., & ten Brinke, L. (2009). Dangerous decisions: A theoretical framework for understanding how judges assess credibility in the courtroom. *Legal and Criminological Psychology*, *14*(1), 119–134. https://doi.org/10.1348/135532508X281520

Porter, S., ten Brinke, L., & Gustaw, C. (2010). Dangerous decisions: The impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime and Law*, *16*(6), 477–491. https://doi.org/10.1080/10683160902926141

Rakison, D. H., & Derringer, J. (2008). Do infants possess an evolved spider-detection mechanism? *Cognition*. https://doi.org/10.1016/j.cognition.2007.07.022

Ramsey, R., & Ward, R. (2020). Challenges and opportunities for top-down modulation research in cognitive psychology. In *Acta Psychologica* (Vol. 209, p. 103118). Elsevier B.V. https://doi.org/10.1016/j.actpsy.2020.103118

Re, D. E., & Rule, N. (2016). Predicting Firm Success From the Facial Appearance of Chief Executive Officers of Non-Profit Organizations. *Perception*, *45*(10), 1137–1150. https://doi.org/10.1177/0301006616652043

Re, D. E., & Rule, N. (2017). Distinctive Facial Cues Predict Leadership Rank and Selection. *Personality and Social Psychology Bulletin*, *43*(9), 1311–1322. https://doi.org/10.1177/0146167217712989

Rees, K., Nadig, A., & Rvachew, S. (2017). Story-related discourse by parent–child dyads: A comparison of typically developing children and children with language impairments. *International Journal of Child-Computer Interaction*, *12*(12), 16–23. https://doi.org/10.1016/j.ijcci.2017.01.001

Repacholi, B. M., Meltzoff, A. N., Toub, T. S., & Ruba, A. L. (2016). *Infants' Generalizations About Other People's Emotions: Foundations for Trait-Like Attributions*. https://doi.org/10.1037/dev0000097

Reyes-García, V., Gallois, S., & Demps, K. (2016). *A Multistage Learning Model for Cultural Transmission: Evidence from Three Indigenous Societies* (pp. 47–60). https://doi.org/10.1007/978-4-431-55997-9_4

Rezlescu, C., Penton, T., Walsh, V., Tsujimura, H., Scott, S. K., & Banissy, M. J. (2015). Dominant Voices and Attractive Faces: The Contribution of Visual and Auditory Information to Integrated Person Impressions. *Journal of Nonverbal Behavior*, *39*(4), 355–370. https://doi.org/10.1007/S10919-015-0214-8

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.*, *57*, 199–226.

Rhodes, M., Leslie, S.-J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences*, *109*(34), 13526–13531. https://doi.org/10.1073/pnas.1208951109

Richter, N., Tiddeman, B., & Haun, D. B. M. (2016). Social preference in preschoolers: Effects of morphological self-similarity and familiarity. *PLoS ONE*, *11*(1). https://doi.org/10.1371/journal.pone.0145443

Rideout, V., & Robb, M. B. (2020). The Common Sense census: Media use by kids age zero to eight, 2020. In *Perception and the Internal Senses*. https://d2e111jq13me73.cloudfront.net/sites/default/files/uploads/research/2020_youngkidsyoutube-report_final-release_forweb.pdf

Ritchie, K. L., Palermo, R., & Rhodes, G. (2017). Forming impressions of facial attractiveness is mandatory. *Scientific Reports*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-00526-9

Rosa, J. “. (2017). Discourse, Linguistic Production, and Subjectivity: Disney-fying Language. *Taboo: The Journal of Culture and Education, 10* (2). https://doi.org/10.31390/taboo.10.2.14

Ross, K. M., Pye, R. E., & Randell, J. (2016). Reading touch screen storybooks with mothers negatively affects 7-year-old readers’ comprehension but enriches emotional engagement. *Frontiers in Psychology*, *7*(NOV), 1728. https://doi.org/10.3389/fpsyg.2016.01728

Roth, Y., & Pickles, N. (2020). *Updating our approach to misleading information*. Twitter Blog. https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information

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

Rule, N. O., & Ambady, N. (2008b). Brief exposures: Male sexual orientation is accurately perceived at 50 ms. *Journal of Experimental Social Psychology*, *44*(4), 1100–1105. https://doi.org/10.1016/J.JESP.2007.12.001

Rule, N. O., & Ambady, N. (2009). She’s Got the Look: Inferences from Female Chief Executive Officers’ Faces Predict their Success. *Sex Roles*, *61*(9–10), 644–652. https://doi.org/10.1007/s11199-009-9658-9

Rule, N. O., & Ambady, N. (2010). Democrats and Republicans Can Be Differentiated from Their Faces. *PLOS ONE*, *5*(1), e8733.

Rule, N. O., Ambady, N., & Hallett, K. C. (2009). Female sexual orientation is perceived accurately, rapidly, and automatically from the face and its features. *Journal of Experimental Social Psychology*, *45*(6), 1245–1251. https://doi.org/10.1016/J.JESP.2009.07.010

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

Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of*

Personality and Social Psychology, *104*(3), 409–426. https://doi.org/10.1037/a0031050

Ryan, M. P., Reese, V., & Wagner, R. F. (2018). Dermatological depictions in animated movies. *British Journal of Dermatology*, *179*(4), 982–983. https://doi.org/10.1111/bjd.16880

Said, C. P., Haxby, J. v., & Todorov, A. (2011). Brain systems for assessing the affective value of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1660–1670. https://doi.org/10.1098/rstb.2010.0351

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

Sakuta, Y., Kanazawa, S., & Yamaguchi, M. K. (2018). Infants prefer a trustworthy person: An early sign of social cognition in infants. *PLOS ONE*, *13*(9), e0203541. https://doi.org/10.1371/journal.pone.0203541

Sassenberg, K., & Ditrich, L. (2019). Research in social psychology changed between 2011 and 2016: Larger sample sizes, more self-report measures, and more online studies. *Advances in Methods and Practices in Psychological Science*, *2*(2), 107-114.

Schaller, M. (2008). Evolutionary bases of first impressions. In N. Ambady & J. J. Skowronski (Eds.), First impressions (pp. 15- 34). New York, NY: Guilford.

Schild, C., Stern, J., & Zettler, I. (2020). Linking men's voice pitch to actual and perceived trustworthiness across domains. *Behavioral Ecology*, *31*(1), 164–175. https://doi.org/10.1093/BEHECO/ARZ173

Segall, G., Birnbaum, D., Deeb, I., & Diesendruck, G. (2015). The intergenerational transmission of ethnic essentialism: how parents talk counts the most. *Developmental Science*, *18*(4), 543–555. https://doi.org/10.1111/DESC.12235

Shahaeian, A., Nielsen, M., Peterson, C. C., & Slaughter, V. (2014). Iranian Mothers' Disciplinary Strategies and Theory of Mind in Children: A Focus on Belief Understanding. *Journal of Cross-Cultural Psychology*, *45*(7), 1110–1123. https://doi.org/10.1177/0022022114534772

Shahi, G. K., Dirkson, A., & Majchrzak, T. A. (2020). An Exploratory Study of COVID-19 Misinformation on Twitter. In *arXiv* (Vol. 22, p. 100104). arXiv. https://doi.org/10.1016/j.osnem.2020.100104

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

Shen, X., Mann, T. C., & Ferguson, M. J. (2020). Beware a dishonest face?: Updating face-based implicit impressions using diagnostic behavioral information. *Journal of Experimental Social Psychology*, *86*. https://doi.org/10.1016/j.jesp.2019.103888

Shimizu, Y., Lee, H., & Uleman, J. S. (2017). Culture as automatic processes for making meaning: Spontaneous trait inferences. *Journal of Experimental Social Psychology*. https://doi.org/10.1016/j.jesp.2016.08.003

Short, L. A., Hatry, A. J., & Mondloch, C. J. (2011). The development of norm-based coding and race-specific face prototypes: An examination of 5- and 8-year-olds' face space. *Journal of Experimental Child Psychology*, *108*(2), 338–357. https://doi.org/10.1016/J.JECP.2010.07.007

Short, L. A., Lee, K., Fu, G., & Mondloch, C. J. (2014). Category-specific face prototypes are emerging, but not yet mature, in 5-year-old children. *Journal of Experimental Child Psychology*, *126*, 161–177. https://doi.org/10.1016/j.jecp.2014.04.004

Shortland, M. (1986). The Power of a Thousand Eyes: Johann Caspar Lavater's Science of Physiognomical Perception. *Criticism: A Quarterly for Literature and the Arts*, *28*(4), 379–408. https://doi.org/10.2307/23110469

Shutts, K., Banaji, M. R., & Spelke, E. S. (2010). Social categories guide young children's preferences for novel objects. *Developmental Science*, *13*(4), 599–610. https://doi.org/10.1111/j.1467-7687.2009.00913.x

Shutts, K., Kinzler, K. D., Katz, R. C., Tredoux, C., & Spelke, E. S. (2011). Race preferences in children: insights from South Africa. *Developmental Science*, *14*(6), 1283–1291. https://doi.org/10.1111/j.1467-7687.2011.01072.x

Siddique, S., Sutherland, C. A., Palermo, R., Foo, Y. Z., Swe, D. C., & Jeffery, L. (2022). Development of face-based trustworthiness impressions in childhood: A systematic review and metaanalysis. *Cognitive Development*, *61*, 101131.

Skinner, A. L., Meltzoff, A. N., & Olson, K. R. (2017). "Catching" social bias: Exposure to biased nonverbal signals creates social biases in preschool children. *Psychological Science*, *28*(2), 216–224. https://doi.org/10.1177/0956797616678930

Skinner, A. L., Olson, K. R., & Meltzoff, A. N. (2019). Acquiring group bias: Observing other people's nonverbal signals can create social group biases. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspi0000218

Slepian, M. L., & Ames, D. R. (2016). Internalized Impressions: The Link Between Apparent Facial Trustworthiness and Deceptive Behavior Is Mediated by Targets' Expectations of How They Will Be Judged. *Psychological Science*. https://doi.org/10.1177/0956797615594897

Sofer, C., Dotsch, R., Oikawa, M., Oikawa, H., Wigboldus, D. H. J., & Todorov, A. (2017). For Your Local Eyes Only: Culture-Specific Face Typicality Influences Perceptions of Trustworthiness. *Perception*, *46*(8), 914–928. https://doi.org/10.1177/0301006617691786

Song, A., & Vul, E. (2020). *Do you see what I see ? A Cross-cultural Comparison of Social Impressions of Faces*. 1714–1720.

Spisak, B. R., Homan, A. C., Grabo, A., & van Vugt, M. (2012). Facing the situation: Testing a biosocial contingency model of leadership in intergroup relations using masculine and feminine faces. *The Leadership Quarterly*, *23*(2), 273–280. https://doi.org/10.1016/j.leaqua.2011.08.006

Stajduhar, A., Ganel, T., Avidan, G., Rosenbaum, R. S., & Freud, E. (2022). Face masks disrupt holistic processing and face perception in school-age children. *Cognitive research: principles and implications*, *7*(1), 1-10.

Stoker, J. I., Garretsen, H., & Spreeuwers, L. J. (2016). The Facial Appearance of CEOs: Faces Signal Selection but Not Performance. *PLOS ONE*, *11*(7), e0159950. https://doi.org/10.1371/journal.pone.0159950

Stolier, R. M., Hehman, E., & Freeman, J. B. (2018). A Dynamic Structure of Social Trait Space. *Trends in Cognitive Sciences*, *22*(3), 197–200. https://doi.org/10.1016/j.tics.2017.12.003

Stolier, R. M., Hehman, E., & Freeman, J. B. (2020). Trait knowledge forms a common structure across social cognition. *Nature Human Behaviour 2020 4:4*, *4*(4), 361–371. https://doi.org/10.1038/s41562-019-0800-6

Stolier, R. M., Hehman, E., Keller, M. D., Walker, M., & Freeman, J. B. (2018). *The conceptual structure of face impressions*. *115*(37), 9210–9215. https://doi.org/10.1073/pnas.1807222115

Stone, A., & Fisher, V. (2020). Changing Negative Perceptions of Individuals With Facial Disfigurement: The Effectiveness of a Brief Intervention. *Basic and Applied Social Psychology*, *42*(5), 341–353. https://doi.org/10.1080/01973533.2020.1768394

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Strouse, G. A., & Samson, J. E. (2021). Learning From Video: A Meta-Analysis of the Video Deficit in Children Ages 0 to 6 Years. *Child Development*, *92*(1), e20–e38. https://doi.org/10.1111/CDEV.13429

Sugden, N. A., Mohamed-Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: Typical, daily exposure to faces documented from a first-person infant perspective. *Developmental Psychobiology*, *56*(2), 249–261.

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

Sutherland, C. A. M., Burton, N. S., Wilmer, J. B., Blokland, G. A. M., Germine, L., Palermo, R., Collova, J. R., & Rhodes, G. (2020a). Individual differences in trust evaluations are shaped mostly by environments, not genes. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(19), 10218–10224. https://doi.org/10.1073/pnas.1920131117

Sutherland, C. A., Collova, J. R., Palermo, R., Germine, L., Rhodes, G., Blokland, G. A., ... & Wilmer, J. B. (2020c). Reply to Cook and Over: Social learning and evolutionary mechanisms are not mutually exclusive. *Proceedings of the National Academy of Sciences*, *117*(28), 16114-16115.

Sutherland, C. A. M., Liu, X., Zhang, L., Chu, Y., Oldmeadow, J. A., & Young, A. W. (2018). Facial First Impressions Across Culture: Data-Driven Modeling of Chinese and British Perceivers' Unconstrained Facial Impressions. *Personality and Social Psychology Bulletin*, *44*(4), 521–537. https://doi.org/10.1177/0146167217744194

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

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

Taggart, J., Ellwood, M., Vasc, D., Chin, S., & Lillard, A. S. (2020). Parents' roles and question-asking during pretend and real activities. *Social Development*, sode.12436. https://doi.org/10.1111/sode.12436

Tang, Y., Harris, P. L., Zou, H., & Xu, Q. (2019). The impact of emotional expressions on children's trust judgments. *Cognition and Emotion*, *33*(2), 318–331. https://doi.org/10.1080/02699931.2018.1449735

Tasimi, A., Dominguez, A., & Wynn, K. (2015). Do-gooder derogation in children: the social costs of generosity. *Frontiers in Psychology*, *0*, 1036. https://doi.org/10.3389/FPSYG.2015.01036

Taylor, M. J., Batty, M., & Itier, R. J. (2004). The faces of development: A review of early face processing over childhood. In *Journal of Cognitive Neuroscience* (Vol. 16, Issue 8, pp. 1426–1442). MIT Press 238 Main St., Suite 500, Cambridge, MA 02142-1046 USA journals-info@mit.edu . https://doi.org/10.1162/0898929042304732

Thierry, S. M., Twele, A. C., & Mondloch, C. J. (2021). Mandatory First Impressions: Happy Expressions Increase Trustworthiness Ratings of Subsequent Neutral Images. *Perception*, *50*(2), 103–115. https://doi.org/10.1177/0301006620987205

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

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, *21*(5), 42–50. https://doi.org/10.1109/38.946630

Tiddeman, B. P., Stirrat, M. R., & Perrett, D. I. (2005). Towards Realism in Facial Image Transformation: Results of a Wavelet MRF Method. *Computer Graphics Forum*, *24*(3), 449–456. https://doi.org/10.1111/j.1467-8659.2005.00870.x

Tigue, C. C., Borak, D. J., O'connor, J. J. M., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, *33*, 210–216. https://doi.org/10.1016/j.evolhumbehav.2011.09.004

Todorov, Alexander., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.83.5.1051

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

Todorov, A., & Oh, D. W. (2021). The structure and perceptual basis of social judgments from faces. *Advances in Experimental Social Psychology*, *63*, 189–245. https://doi.org/10.1016/BS.AESP.2020.11.004

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

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, *27*(6), 813–833. https://doi.org/10.1521/soco.2009.27.6.813

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

Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology*, *39*(6), 549–562. https://doi.org/10.1016/S0022-1031(03)00059-3

Toet, A., Bijlsma, M., & Brouwer, A.-M. (2017). Stress Response and Facial Trustworthiness Judgments in Civilians and Military. *SAGE Open*, *7*(3), 215824401772538. https://doi.org/10.1177/2158244017725386

Tsantani, M., Podgajecka, V., Gray, K. L., & Cook, R. (2022). How does the presence of a surgical face mask impair the perceived intensity of facial emotions?. *PloS one*, *17*(1), e0262344.

Turton, D. (2004). Lip-plates and "the people who take photographs": Uneasy encounters between Mursi and tourists in southern Ethiopia. *Anthropology Today*, *20*(3). https://doi.org/10.1111/j.0268-540x.2004.00266.x

Uleman, J. S., Newman, L., & Winter, L. (1992). Can personality traits be inferred automatically? Spontaneous inferences require cognitive capacity at encoding. *Consciousness and Cognition*, *1*(1), 77–90. https://doi.org/10.1016/1053-8100(92)90049-G

Valentine, T. (2018). A Unified Account of the Effects of Distinctiveness, Inversion, and Race in Face Recognition, *43*(2), 161–204. https://doi.org/10.1080/14640749108400966

Vanderborght, M., & Jaswal, V. (2009). Who Knows Best? Preschoolers Sometimes Prefer Child Informants Over Adult Informants. *Infant and Child Development*, *18*(1), 61–71. https://doi.org/10.1002/ICD.591

van der Zant, T., Reid, J., Mondloch, C. J., & Nelson, N. L. (2021). The influence of postural emotion cues on implicit trait judgements. *Motivation and Emotion*, *45*(5). https://doi.org/10.1007/s11031-021-09889-z

van Vugt, M., & Grabo, A. E. (2015). The Many Faces of Leadership: An Evolutionary-Psychology Approach. *Current Directions in Psychological Science*, *24*(6), 484–489. https://doi.org/10.1177/0963721415601971

Verosky, S. C., & Todorov, A. (2010). Generalization of Affective Learning About Faces to Perceptually Similar Faces. *Psychological Science*, *21*(6), 779–785. https://doi.org/10.1177/0956797610371965

Verosky, S. C., & Todorov, A. (2013). When physical similarity matters: Mechanisms underlying affective learning generalization to the evaluation of novel faces. *Journal of Experimental Social Psychology*, *49*(4), 661–669. https://doi.org/10.1016/j.jesp.2013.02.004

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Vrij, A. (1997). Wearing Black Clothes: The Impact of Offenders' and Suspects' Clothing on Impression Formation. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition* , *11*(1), 47–53. https://doi.org/10.1002/(SICI)1099-0720(199702)11:1

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

Walle, E. A., Reschke, P. J., & Knothe, J. M. (2017). Social Referencing: Defining and Delineating a Basic Process of Emotion. *Emotion Review*, *9*(3), 245–252. https://doi.org/10.1177/1754073916669594

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

Wang, M., Yan, B., Yang, F., & Zhao, Y. (2018). The development of spontaneous trait inferences about the actor and spontaneous trait transferences about the informant: Evidence from children aged 8-13 years. *International Journal of Psychology*, *53*(4), 269–277. https://doi.org/10.1002/ijop.12367

Weisbuch, M., Pauker, K., & Ambady, N. (2009). The subtle transmission of race bias via televised nonverbal behavior. *Science (New York, N.Y.)*, *326*(5960), 1711–1714. https://doi.org/10.1126/science.1178358

Willard, G., Isaac, K. J., & Carney, D. R. (2015). Some evidence for the nonverbal contagion of racial bias. *Organizational Behavior and Human Decision Processes*, *128*, 96–107. https://doi.org/10.1016/j.obhdp.2015.04.002

Willis, J., & Todorov, A. (2006). First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science*, *17*(7), 592–598. https://doi.org/10.1111/j.1467-9280.2006.01750.x

Willis, M. L., Palermo, R., & Burke, D. (2011). Social judgments are influenced by both facial expression and direction of eye gaze. *Social Cognition*, *29*(4), 415–429.

Wilson, J. P., & Rule, N. O. (2015). Facial Trustworthiness Predicts Extreme Criminal-Sentencing Outcomes. *Psychological Science*, *26*(8), 1325–1331. https://doi.org/10.1177/0956797615590992

Wong, S. H., & Zeng, Y. (2017). Do inferences of competence from faces predict political selection in authoritarian regimes? Evidence from China. *Social Science Research*, *66*, 248–263. https://doi.org/10.1016/j.ssresearch.2016.11.002

Workman, C. I., Humphries, S., Hartung, F., Aguirre, G. K., Kable, J. W., & Chatterjee, A. (2021). Morality is in the eye of the beholder: the neurocognitive basis of the “anomalous-is-bad” stereotype. *Annals of the New York Academy of Sciences*, *1494*(1), 3–17. https://doi.org/10.1111/NYAS.14575

Xiao, N. G., Quinn, P. C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2018). Older but not younger infants associate own-race faces with happy music and other-race faces with sad music. *Developmental Science*, *21*(2), e12537. https://doi.org/10.1111/desc.12537

Xie, S. Y., Flake, J. K., Stolier, R. M., Freeman, J. B., & Hehman, E. (2021). Facial impressions are predicted by the structure of group stereotypes. *Psychological Science.*

Zebrowitz, L. A. (2004). The Origin of First Impressions. *Journal of Cultural and Evolutionary Psychology*, *2*(1), 93–108. https://doi.org/10.1556/JCEP.2.2004.1-2.6

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

Zebrowitz, L. A., Bronstad, P. M., & Lee, H. K. (2007). The Contribution of Face Familiarity to Ingroup Favoritism and Stereotyping. *Social Cognition*, *25*(2), 306–338. https://doi.org/10.1521/soco.2007.25.2.306

Zebrowitz, L. A., Collins, M. A., & Dutta, R. (1998). The relationship between appearance and personality across the life span. *Personality and Social Psychology Bulletin*, *24*(7), 736–749. https://doi.org/10.1177/0146167298247006

Zebrowitz, L. A., Fellous, J. M., Mignault, A., & Andreoletti, C. (2003). Trait impressions as overgeneralized responses to adaptively significant facial qualities: Evidence from connectionist modeling. *Personality and Social Psychology Review*, *7*(3), 194–215. https://doi.org/10.1207/S15327957PSPR0703_01

Zebrowitz, L. A., Kikuchi, M., & Fellous, J.-M. (2010). Facial resemblance to emotions: group differences, impression effects, and race stereotypes. *Journal of Personality and Social Psychology*, *98*(2), 175.

Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*, *15*(6), 603–623. https://doi.org/10.1007/BF01065855

Zebrowitz, L. A., & Montepare, J. M. (2008). Social Psychological Face Perception: Why Appearance Matters. *Social and Personality Psychology Compass*, *2*(3), 1497. https://doi.org/10.1111/J.1751-9004.2008.00109.X

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

Zebrowitz, L. A., & Zhang, Y. (2009). The Origins of First Impressions in Animal and Infant Face Perception. In *New York* (Vol. 389, Issue September). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195342161.013.0029

Zebrowitz, L. A., & Zhang, Y. (2011). *Origins of impression formation in animal and infant face perception*. The handbook of social neuroscience. Oxford: Oxford University Press.

Zhang, Q., & Wang, M. (2013). The Development of Spontaneous Trait Inferences: Evidence from Chinese Children: *Http://Dx.Doi.Org/10.2466/21.07.PR0.112.3.887-899*, *112*(3), 887–899. https://doi.org/10.2466/21.07.PR0.112.3.887-899

Zhou, X., Matthews, C. M., Baker, K. A., & Mondloch, C. J. (2018). Becoming Familiar With a Newly Encountered Face: Evidence of an Own-Race Advantage. *Perception*, *47*(8), 807–820. https://doi.org/10.1177/0301006618783915

Zucker, J. K., & Patterson, M. M. (2018). Racial Socialization Practices Among White American Parents: Relations to Racial Attitudes, Racial Identity, and School Diversity. *Journal of Family Issues*, *39*(16), 3903–3930. https://doi.org/10.1177/0192513X18800766

Appendices

Appendix A

Chapter 3 – Target pairs average pre-ratings of trustworthiness/niceness

Table A

Average (SD) trustworthiness/niceness ratings for targets by pairings.

Gender	Target A	Target B
Male	36.80 (20.44)	37.30 (22.96)
Male	31.45 (23.76)	35.10 (19.98)
Male	54.30 (20.62)	55.10 (19.00)
Male	57.10 (14.84)	59.60 (14.35)
Female	30.90 (20.23)	36.00 (20.50)
Female	42.00 (18.48)	48.90 (17.35)
Female	49.05 (19.45)	49.15 (17.45)
Female	55.55 (15.00)	57.8 (16.89)

Appendix B

Chapter 4 - Study 1: Mixed Models. Tables A-C

Table A

Study 1: Linear mixed models fixed and random effects comparing the numbers of words spoken by parents between conditions (left) and children between conditions (right).

	Parent Word Count			Child Word Count		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
Intercept (Low Trustworthiness)	78.04	57.20 – 98.88	**<0.001**	34.42	23.68 – 45.15	**<0.001**
High Competence	5.13	-7.12 – 17.37	0.412	-9.42	-17.87 – -0.96	**0.029**
High Trustworthiness	1.71	-10.53 – 13.95	0.784	-0.63	-9.08 – 7.83	0.885
Low Competence	0.46	-11.78 – 12.70	0.942	2.75	-5.70 – 11.20	0.524
Random Effects						
σ^2	468.21			223.24		
τ_{00}	2244.58 participant_id			497.10 participant_id		
ICC	0.83			0.69		
N	24 participant_id			24 participant_id		
Observations	96			96		
Marginal R^2 / Conditional R^2	0.001 / 0.828			0.028 / 0.699		

Table B
Study 1: Generalised linear mixed effect models, odds ratios and random effects comparing likelihood of parents (left) and children (right) using trait terms between conditions.

	Parent Traits – Binomial			**Child Traits - Binomial**		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
Intercept (Low Trustworthiness)	2.47	0.38 – 16.23	0.346	1.41	0.27 – 7.36	0.681
High Competence	0.32	0.05 – 1.87	0.205	0.35	0.07 – 1.91	0.227
High Trustworthiness	0.47	0.08 – 2.65	0.391	0.35	0.07 – 1.91	0.227
Low Competence	0.21	.03 – 1.34	0.099	0.17	0.03 – 1.03	0.054
Random Effects						
σ^2	3.29			3.29		
τ_{00}	10.58 $_{participant\_id}$			7.84 $_{participant\_id}$		
ICC	0.76			0.70		
N	24 $_{participant\_id}$			24 $_{participant\_id}$		
Observations	96			96		
Marginal R^2 / Conditional R^2	0.023 / 0.768			0.035 / 0.715		

Table C
Study 1: Generalised linear mixed effect models, odds ratios and random effects comparing likelihood of parents (left) and children (right) using emotion terms between conditions.

	Parent Emotions – Binomial			Child Emotions - Binomial		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
Intercept (Low Trustworthiness)	0.77	0.20 – 3.02	0.708	0.46	0.13 – 1.61	0.224
High Competence	0.75	0.17 – 3.34	0.704	0.30	0.06 – 1.47	0.138
High Trustworthiness	0.75	0.17 – 3.34	0.704	2.17	0.52 – 9.14	0.290
Low Competence	0.41	0.09 – 1.92	0.257	0.57	0.13 – 2.51	0.459
Random Effects						
σ^2	3.29			3.29		
τ_{00}	4.43 $_{participant\_id}$			2.92 $_{participant\_id}$		
ICC	0.57			0.47		
N	24 $_{participant\_id}$			24 $_{participant\_id}$		
Observations	96			96		
Marginal R^2 / Conditional R^2	0.014 / 0.580			0.079 / 0.512		

Appendix C

Chapter 4 - Study 1 Trait and Emotion Terms: Tables A-D

Table A

Parent Trait Terms

Face	Term	Frequency
High Trustworthiness	Nice	9
	Kind	5
	Friendly	3
	Friendlier	2
	Naughty	2
	Sporty	2
	Good	1
	Grumpy	1
	Helpful	1
	Mean	1
Low Trustworthiness	Boring	6
	Nice	6
	Friendly	4
	Bad	3
	Good	3
	Grumpy	3
	Kind	3
	Meanie	2
	Scary	2
	Creative	1
	Friendlier	1
	Intimidating	1
	Nasty	1
	Naughty	1
	Sensible	1
	Sporty	1
High Competence	Nice	10
	Adventurous	3
	Brave	3
	Cheerful	3
	Good	3
	Intelligent	3
	Kind	3
	Mean	3
	Friendly	2
	Grumpy	2
	Nasty	2

Table A - Continued

High Competence	Caring	1
	Helpful	1
	Scary	1
	Sporty	1
Low Competence	Nice	7
	Kind	5
	Evil	4
	Friendly	4
	Good	2
	Grumpy	2
	Scary	2
	Depressed	1
	Depression	1
	Mean	1
	Nosey	1

Table B

Child Trait Terms

Face	Term	Frequency
High Trustworthiness	Nice	4
	Kind	2
	Friendly	1
	Good	1
	Naughty	1
	Sporty	1
Low Trustworthiness	Boring	4
	Smart	3
	Bad	1
	Good	1
	Grump	1
	Grumpy	1
	Meanie	1
	Nasty	1
	Naughty	1
	Nice	1
	Scary	1
	Sensible	1
	Sporty	1
High Competence	Nice	4
	Adventurous	1
	Brave	1
	Cheerful	1
	Good	1
	Grumpy	1
	Mean	1
	Scary	1
Low Competence	Depressed	2
	Evil	2
	Scary	2
	Bad	1
	Good	1
	Grumpy	1
	Kind	1
	Nice	1
	Nosey	1

Table C

Parent Emotion & Expression Terms

Face	Term	Frequency
High Trustworthiness	Happy	15
	Sad	5
	Smile	3
	Smiling	2
	Angry	1
	Happier	1
	Straight Face	1
Low Trustworthiness	Happy	10
	Sad	6
	Angry	5
	Moody	2
	Cross	1
	Frowning	1
	Joyful	1
High Competence	Happy	8
	Sad	6
	Crosser	3
	Smiling	3
	Angry	1
	Happier	1
	Smile	1
	Smiley	1
	Straight face	1
Low Competence	Happy	7
	Sad	6
	Cross	1
	Nervous	1
	Sadder	1
	Scared	1
	Smiling	1
	Surprised	1
	Tearful	1

Table D

Child Emotion & Expression Terms

Face	Term	Frequency
High Trustworthiness	Happy	9
	Cross	1
	Happier	1
	Smile	1
	Smiley	1
	Smiling	1
	Straight Face	1
Low Trustworthiness	Angry	4
	Sad	3
	Cross	1
	Frowning	1
	Joyful	1
	Moody	1
High Competence	Crosser	2
	Cross	1
	Happy	1
	Smile	1
	Smiling	1
	Straight Face	1
Low Competence	Sad	3
	Happy	2
	Frowny	1
	Sadder	1
	Smiling	1

Appendix D

Chapter 4 - Study 2: Mixed Models. Tables A-C

Table A

Study 2: Linear mixed models fixed and random effects comparing the numbers of words spoken by parents between conditions (left) and children between conditions (right)..

	Parent Word Count			Child Word Count		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept) Low Trustworthiness	76.62	56.99 – 96.26	**<0.001**	25.62	19.64 – 31.61	**<0.001**
High Intelligence	4.79	-12.02 – 21.60	0.576	-2.17	-7.74 – 3.40	0.446
High Trustworthiness	-3.46	-20.27 – 13.35	0.687	-2.58	-8.15 – 2.99	0.363
Low Intelligence	-7.29	-24.10 – 9.52	0.395	-2.42	-7.99 – 3.15	0.395
Random Effects						
σ^2	882.56			96.90		
τ_{00}	1526.20 participant_id			127.08 participant_id		
ICC	0.63			0.57		
N	24 participant_id			24 participant_id		
Observations	96			96		
Marginal R^2 / Conditional R^2	0.008 / 0.637			0.005 / 0.569		

Table B
Study 2: Generalised linear mixed effect models, odds ratios and random effects comparing likelihood of parents (left) and children (right) using trait terms between conditions.

	Parent Traits – Binomial			Child Traits - Binomial		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept) Low Trustworthiness	1.27	0.41 – 3.97	0.683	1.25	0.41 – 3.77	0.697
High Intelligence	0.23	0.05 – 0.98	**0.047**	0.31	0.08 – 1.25	0.099
High Trustworthiness	0.49	0.12 – 1.93	0.308	0.39	0.10 – 1.55	0.183
Low Intelligence	1.00	0.26 – 3.87	1.000	0.63	0.16 – 2.41	0.500
Random Effects						
σ^2	3.29			3.29		
τ_{00}	2.29 participant_id			2.01 participant_id		
ICC	0.41			0.38		
N	24 participant_id			24 participant_id		
Observations	96			96		
Marginal R^2 / Conditional R^2	0.064 / 0.448			0.038 / 0.403		

Table C
Study 2: Generalised linear mixed effect models, odds ratios and random effects comparing likelihood of parents (left) and children (right) using emotion terms between conditions.

	Parent Emotions – Binomial			Child Emotions - Binomial		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
Intercept (Low Trustworthiness)	0.67	0.10 – 4.31	0.670	0.65	0.09 – 4.83	0.670
High Intelligence	7.62	1.04 – 55.67	**0.045**	3.57	0.54 – 23.87	0.189
High Trustworthiness	19.11	2.05 – 177.77	**0.010**	23.51	2.17 – 254.63	**0.009**
Low Intelligence	3.30	0.53 – 20.62	0.202	1.00	0.16 – 6.06	1.000
Random Effects						
σ^2	3.29			3.29		
τ_{00}	10.85 $_{participant\_id}$			13.07 $_{participant\_id}$		
ICC	0.77			0.80		
N	24 $_{participant\_id}$			24 $_{participant\_id}$		
Observations	96			96		
Marginal R^2 / Conditional R^2	0.078 / 0.785			0.094 / 0.818		

Appendix E

Chapter 4 - Study 2 Trait and Emotion Terms: Tables A-D

Table A

Parent Trait Terms

Face	Term	Frequency
High Trustworthiness	Nice	5
	Serious	4
	Cool	2
	Goody	2
	Scary	2
	Awesome	1
	Bad	1
	Crazy	1
	Good	1
	Kind	1
	Nasty	1
Low Trustworthiness	Scary	9
	Naughty	6
	Baddy	3
	Mean	3
	Clever	2
	Good	2
	Goody	2
	Mad	2
	Nice	2
	Cleverer	1
	Grumpy	1
	Lazy	1
High Intelligence	Nice	5
	Bad	4
	Friendly	4
	Scary	3
	Good	2
	Kind	2
	Naughty	2
	Clever	1
	Crazy	1
	Serious	1
	Shifty	1
Low Intelligence	Crazy	8
	Grumpy	4
	Good	3
	Lazy	3
	Goody	2

Table A - Continued

Low Intelligence	Silly	2
	Bad	1
	Mad	1
	Naughty	1
	Scary	1
	Serious	1
	Trustworthy	1
	Untrustworthy	1

Table B

Child Trait Terms

Face	Term	Frequency
High Trustworthiness	Serious	3
	Awesome	2
	Nice	2
	Bad	1
	Cool	1
	Goody	1
Low Trustworthiness	Goody	2
	Scary	2
	Bad	1
	Baddy	1
	Clever	1
	Daring	1
	Grumpy	1
	Lazy	1
	Mad	1
	Mean	1
	Naughty	1
	Suspicious	1
High Intelligence	Nice	2
	Bad	1
	Good	1
	Kind	1
	Naughty	1
	Scary	1
Low Intelligence	Grumpy	6
	Crazy	3
	Good	2
	Lazy	2
	Goody	1
	Mad	1
	Naughty	1
	Scary	1
	Silly	1

Table C

Parent Emotion & Expression Terms

Face	Term	Frequency
High Trustworthiness	Happy	17
	Sad	16
	Bored	5
	Confused	3
	Smile	3
	Angry	1
	Frown	1
	Miserable	1
	Smiles	1
	Smiling	1
	Unhappy	1
Low Trustworthiness	Sad	13
	Happy	11
	Angry	10
	Bored	2
	Smiling	2
	Unhappy	2
	Confused	1
	Cross	1
High Intelligence	Happy	15
	Sad	8
	Smile	5
	Confused	4
	Happier	4
	Angry	3
	Smiling	2
	Scowls	1
	Straight Faced	1
Low Intelligence	Sad	22
	Happy	12
	Confused	3
	Angry	2
	Cross	2
	Nervous	2
	Unsure	1

Table D

Child Emotion & Expression Terms

Face	Term	Frequency
High Trustworthiness	Sad	7
	Happy	6
	Bored	3
	Confused	3
	Frown	1
	Miserable	1
Low Trustworthiness	Sad	5
	Angry	4
	Bored	1
	Confused	1
	Cross	1
	Happier	1
	Happy	1
	Sadder	1
	Smiling	1
High Intelligence	Happy	8
	Confused	2
	Sad	2
	Smile	2
	Angry	1
Low Intelligence	Sad	6
	Happy	2
	Nervous	2
	Angry	1
	Confused	1
	Cross	1
	Smile	1