

**First impressions of child faces: Facial trustworthiness influences adults' interpretations of
children's behaviour in ambiguous situations**

by

Sophia Thierry

Department of Psychology

Brock University

St. Catharines, Ontario

© Sophia Thierry, 2021

Abstract

Despite the profound behavioural consequences that first impressions of trustworthiness have on adult populations, few studies have examined how adults' first impressions of trustworthiness influence behavioural outcomes for children. Using a novel task design, we examined adults' perceptions of children's behaviour in ambiguous situations. After a brief presentation of a child's face (high or low trust), participants viewed the child's face embedded within an ambiguous scene involving two children (Scene Task) or read a vignette about a misbehaviour done by that child (Misbehaviour Task). In the Scene Task, participants described what they believed to be happening in each scene; in the Misbehaviour Task, participants indicated whether the behaviour was done on purpose or by accident. In both tasks, participants also rated the behaviour of the target child and indicated whether that child would be a good friend. In Experiment 1, young ($n=61$) and older ($n=57$) adults viewed unaltered face images. Ambiguous scenes and misbehaviours were interpreted more positively when the target child had a high- versus low-trust face, with comparable patterns of results for the two age groups. In Experiment 2, young adults ($N=59$) completed the same tasks while viewing images of child faces morphed towards high- and low-trust averages. The pattern of results mirrored that of Experiment 1. Collectively, our results demonstrate that a child's facial trustworthiness biases how adults interpret children's behaviour—a heuristic that may have lasting behavioural consequences for children through a self-fulfilling prophecy.

Acknowledgments

I would like to thank my supervisor, Dr. Catherine Mondloch for her unwavering support, kindness and encouragement. Thank you for inspiring me each and every day to grow as a researcher and as a student. I am deeply appreciative of your guidance.

I would like to thank my committee members, Dr. Angela Evans and Dr. Caitlin Mahy for your mentorships throughout my master's degree. You have both inspired my research and course work. I am sincerely grateful to have had such encouraging committee members.

Finally, I would like to acknowledge everyone who played a role in my academic accomplishments. I appreciate the support, encouragement and friendship of my family, lab-mates, and friends.

Table of Contents

ABSTRACT	II
ACKNOWLEDGEMENTS.....	III
LIST OF FIGURES	VI
LIST OF TABLES	VII
INTRODUCTION.....	1
EXPERIMENT 1.....	6
METHOD.....	6
PARTICIPANTS.....	8
GENERAL PROCEDURE.....	7
FACE STIMULI.....	8
SCENE TASK.....	8
STIMULI	8
PROCEDURE	8
CODING PROCEDURE.....	9
MISBEHAVIOUR TASK.....	12
STIMULI	12
PROCEDURE	12
TRAIT RECOGNITION TASK.....	14
STIMULI	14
PROCEDURE	14
RESULTS.....	14
ANALYTIC APPROACH.....	14
TRAIT RECOGNITION TASK.....	15
SCENE TASK.....	15
INTERPRETATION	15
BEHAVIOURAL RATING.....	16
FRIENDSHIP EVALUATION.....	16
MISBEHAVIOUR TASK.....	16
INTERPRETATION	16
BEHAVIOURAL RATING.....	17
FRIENDSHIP EVALUATION.....	17
DISCUSSION	22
EXPERIMENT 2.....	22
METHODS	23
PARTICIPANTS.....	23

GENERAL PROCEDURE.....	23
FACE STIMULI.....	23
RESULTS.....	24
ANALYTIC APPROACH.....	24
TRAIT RECOGNITION TASK.....	24
SCENE TASK.....	24
INTERPRETATION	24
BEHAVIOURAL RATING.....	25
FRIENDSHIP EVALUATION.....	25
MISBEHAVIOUR TASK.....	25
INTERPRETATION	25
BEHAVIOURAL RATING.....	26
FRIENDSHIP EVALUATION.....	26
DISCUSSION	26
GENERAL DISCUSSION.....	27
REFERENCES.....	34

List of Figures

Figure 1. Example of situations presented in the Scene Task.....	11
Figure 2. Interpretations provided: Scene Task.....	19
Figure 3. Interpretations provided: Misbehaviour Task.....	20

List of Tables

Table 1. Vignettes presented in the Misbehaviour Task.....	13
Table 2. Results of single sample t-test in Experiment 1 and 2.....	18

First impressions of child faces: Facial trustworthiness influences adults' interpretations of children's behaviour in ambiguous situations¹

A first glance at a face is enough to form inferences about a person's character. Indeed, social attributions from faces are made readily within 100ms of seeing a face (Willis & Todorov, 2006). First impressions from faces have serious real-world outcomes for adult populations (*see* Harris & Garris, 2008; Rule & Ambady, 2010; Todorov, 2017; Zebrowitz & Montepare, 2015; Zebrowitz, 2017 *for review*). First impressions of adult faces predict election outcomes (Ballew & Todorov, 2007; Castelli et al., 2009; Olivola & Todorov, 2010a; Todorov et al., 2005), hiring decisions (Centorrino et al., 2015; Rule & Ambady, 2008; 2011), healthcare priority (Bagnis et al., 2020; Mattarozzi et al., 2017) and, even more powerfully, sentencing decisions (Jaeger et al., 2020; Wilson & Rule, 2015).

Surprisingly little is known about adults' first impressions of child faces. Adults are quick to ascribe character traits such as “nice” and “mean” to children's faces varying in trustworthiness (Cogsdill & Banaji, 2015; Cogsdill et al., 2014; Collova et al., 2019), impressions that are at risk of having cascading effects on development. Children perceived as “nice” are likely to be treated more positively than children perceived as “mean” and enjoy more positive interactions with adults as a result—a possibility that has received scant attention in the literature. In the current study, we examine for the first time, how facial trustworthiness influences adults' interpretations of children's behaviour in ambiguous situations. We

¹ This thesis is adapted from the published article: Thierry, S. M., & Mondloch, C.J. (2021). First impressions of child faces: Facial trustworthiness influences adults' interpretations of children's behaviour in ambiguous situations. *Journal of Experimental Child Psychology*, 208, 105153. doi: 10.1016/j.jecp.2021.105153

investigated whether children perceived as “nice” would be perceived more positively than children perceived as “mean” for the same behavioural act.

Adults’ treatment of children is influenced by facial appearance. For example, adults rated children who were smiling in videos as cuter than children who were crying or neutral in expression and indicated greater likelihood of adopting them (Aradhye et al., 2015). Adults judged misbehaviours of children with mature facial features as more intentional, and in some instances deserving of more severe punishment, as compared to their babyface peers (Zebrowitz et al., 1991). Further, adults showed a halo effect for children high in attractiveness, providing more positive evaluations of those children on a variety of characteristics such as intelligence, leadership ability, and sociability (Clifford & Walster, 1973; Kenealy et al., 1988; Langlois et al., 2000). Given this evidence that facial appearance influences adults’ behaviour towards children, it is imperative that we understand the social consequences of first impressions of traits for children as well.

Of all the trait impressions underlying adult faces, trustworthiness is the primary dimension and signals one’s intent to help vs. harm (Oosterhof & Todorov, 2008; Sutherland et al., 2018; 2016). The primacy of trustworthiness aligns with functional accounts suggesting that first impressions based on subtle facial cues are an overgeneralization of an adaptation to detect potential threat (Oosterhof & Todorov, 2008). The only study to examine the dimensions underlying adults’ first impressions of child faces showed that niceness is the primary dimension and is similar to trustworthiness (i.e., adults’ ratings of trustworthiness and niceness correlate strongly for child faces; Collova et al., 2019). Here, we focus on the influence of facial trustworthiness/niceness on adults’ perceptions of children’s intent in ambiguous situations.

In contrast to the extensive literature examining the social consequences of first impressions of trustworthiness in adult populations (*see* Todorov, 2017; Zebrowitz & Montepare, 2015; Zebrowitz, 2017 *for review*), only two recent studies have examined behavioural consequences of adults' perceptions of children's facial trustworthiness. When shown a pair of child faces (one high- and one low-nice), adults selected the high-nice child when asked which child they would be most likely to reward for good behaviour and selected the low-nice child when asked which child they would be most likely to watch more closely for rough play (Collova et al., 2019). In an economic trust paradigm, adults were more trusting of (i.e., invested more tokens in) child and adult partners displaying slightly happy versus slightly angry faces (Ewing & al., 2019); multiple studies have shown that expressions of happiness and anger produce perceptions of high- and low-trust respectively (Oosterhof & Todorov, 2008; Said et al., 2009; Caulfield et al., 2016). Together, these results highlight the power of first impressions in influencing adults' behavioural expectations of children (Collova et al., 2019) and their behaviour towards them (Ewing et al., 2019), a bias that is likely to impact real-life outcomes through a self-fulfilling prophecy (Merton, 1948). The current study builds on these findings to examine the influence of adults' first impressions of child faces in ambiguous situations.

To examine adults' interpretations of children's behaviour in ambiguous situations, we designed the Ambiguous Situations protocol which is comprised of the Scene Task (adapted from McGlothlin et al., 2005) and the Misbehaviour Task (adapted from Zebrowitz, et al, 1991). On each trial participants viewed a child's face that was previously rated as high or low in trustworthiness. In the Scene Task, participants then view the child's face embedded within an ambiguous scene; each scene involves two children, and the actions of the target child could be perceived as pro- or anti-social. In the Misbehaviour Task, participants then read a vignette

describing a misbehaviour done by that child; each vignette is worded such that the child's intentions are unclear (i.e., the behaviour might have been done on purpose or by accident). Participants are asked to interpret each situation and then rate how good each child's behaviour was and whether each child would be a good friend. Whereas in the Scene Task the child's behaviour is ambiguous (e.g., is the target child helping to build the tower or about to knock the tower down?), in the Misbehaviour Task the intention behind a misbehaviour is ambiguous (e.g., did the child spill the paint on purpose or by accident?). If children's facial trustworthiness influences adults' interpretations of their behaviour, the scenes and misbehaviours should be interpreted more positively when the target child has a high- versus a low-trust face.

An advantage of our task is that we use neutral faces as opposed to emotional expressions (*as in* Ewing et al., 2019), and presented faces one at a time, as opposed to in pairs (*as in* Collova et al., 2019) reflecting how first impressions are formed in real-life settings. In addition, the two previous studies examined adults' predictions of children's behaviour (i.e., the likelihood for a child to return an investment, Ewing et al., 2019; and the likelihood for a child to engage in certain behaviours, Collova et al., 2019). In the current study, we are directly examining adults' perceptions of the child's *intention* driving the behaviour at hand. As such, our research is comparable to studies examining the social consequences of adults' evaluations of intention in adult faces. For example, when evaluating sentencing decisions based on hypothetical small claims court cases, defendants were more likely to be found guilty when presented with an untrustworthy (vs. trustworthy) face (Jaeger et al., 2020), showing the influence of first impressions in biasing perceptions of intent in adult populations. Building directly upon the functional account, our study examines adults' interpretations of children's intent to help vs. harm. Ambiguous situations are an ideal context for examining the role of trustworthiness in

signaling intention to help versus harm because the interpretation of the target's behaviour (e.g., whether a misbehaviour was intentional or unintentional) lies in the eye of the beholder.

In Experiment 1 we included both young (ages 18-27) and older adults (ages 65+) to see whether there are age differences in adults' perceptions of children's behaviour. We predicted that relative to younger adults, older adults would rate low-trust faces more positively and show less of a bias when interpreting scenes and misbehaviours in which the target child has a low-trust face. This prediction is based on previous research suggesting that older adults demonstrate a positivity bias for many stimulus categories (Mather & Carstensen, 2005) at least under some task conditions (*see* Zebrowitz et al., 2017). For example, older adults are less sensitive to adult faces low in trustworthiness (i.e., provide more positive ratings; Zebrowitz et al., 2013) and are less sensitive to children's lie telling (i.e., have a truth bias; O'Connor et al., 2019).

In Experiment 2, we replicated findings from Experiment 1 using a more controlled set of facial stimuli to verify that our effects were driven by variations of trustworthiness. Young adults completed the same tasks while viewing images of average-trust child faces that were morphed towards high- and low-trust face averages. Participants saw either the high- or low-trust version of the same identity. The use of facial identities morphed towards a high-, and low-trust face average allowed for a stronger test of causality as we were able to control for facial identity in influencing ratings of trustworthiness. In addition, the morphed identities presented more subtle variations of trustworthiness. Experiment 2 allowed us to investigate whether subtle expressions of trustworthiness biased adults' interpretations of children's behaviour in ambiguous situations, providing an important replication.

Experiment 1

The main purpose of Experiment 1 was to examine the influence of facial trustworthiness on young and older adults' interpretations of children's behaviour in ambiguous situations. Participants viewed a child's face embedded within an ambiguous scene (Scene Task) or prior to reading a vignette about a misbehaviour (Misbehaviour Task); each face had previously been rated as high or low in trustworthiness (Collova et al., 2019). Participants were instructed to tell a story about each ambiguous scene and to indicate whether each misbehaviour (e.g., a child spilling paint on another child's drawing) happened on purpose or by accident. Participants also rated how good each target child's behaviour was and whether each target child would be a good friend. We predicted that the ambiguous scenes would be interpreted as more prosocial, and that a higher proportion of misbehaviours would be rated as accidental, when the target child had a high-trust versus a low-trust face. We predicted that the target child's behaviour would be rated more positively, and that the target child was more likely to be perceived as being a good friend, when the target child had a high-trust than a low-trust face. Such findings would shed new insights on how facial cues might lead to differential treatment by adults.

We predicted that this difference would be less pronounced for older adults who might show a positivity bias (Mather & Carstensen, 2005) and thus might interpret the behaviour of children with low trust faces more positively than young adults.

Method

Participants

Sixty-one young adults (age range: 18-27; $M = 21.98$, $SD = 2.67$; 41 female; $n=51$ Caucasian) were recruited online through Testable Minds (Testable S.R.L., 2018, $n=27$, 67% from the US, 30% from the UK, 3% from Canada) and the community ($n=34$, 97% from Canada,

3% from the US). Fifty-seven older adults (age range: 65-79, $M = 70.30$, $SD = 5.34$; 36 female; $n=56$ Caucasian) were recruited online through Amazon Mechanical Turk (2018; $n=29$, all from the US) and the community ($n=28$, 89% from Canada, 11% from the US). All participants were tested online. Preliminary analyses indicated no significant differences between participants tested online through experimental vs. community platforms, $ps > .200$; therefore, the online data were combined.

Prior to collecting data, we aimed to test 60 young and 60 older adults to establish our novel tasks for future use with children. G*Power 3.1 confirmed that this sample size exceeded that required to find a medium effect size ($F=4.15$) in a 2 (between group) x 2 (repeated measures) ANOVA with power = 0.8. Participants recruited on Testable Minds and Amazon Mechanical Turk were compensated \$2.00 USD; participants recruited in the community received entry into a draw to win one of two \$40.00 CAD gift cards. An additional two participants were excluded for failing attention checks designed to ensure participants understood and complied with task instructions (one young adult; one older adult; see “procedure” section of *Trait Recognition*).

General Procedure

The Ambiguous Situations protocol (i.e., Scene and Misbehaviour Task) was a part of a larger battery of tasks completed online on Qualtrics (Qualtrics, Provo, UT). Participants completed the tasks in this fixed order: Scene, Inferential Reasoning, Misbehaviour, Trait Recognition, and Trait Understanding. The Inferential Reasoning and Trait Understanding Task will be used in future research with children but were not analyzed for this paper as performance was at ceiling. The project received clearance from the Research Ethics Board at Brock University.

Face Stimuli

Photographs of 12 Caucasian identities (six male, six female; ages 4-6), each displaying a neutral expression were selected from the Child Affect Facial Expression set (CAFE; LoBue, 2014; LoBue & Thrasher, 2015). Four identities were assigned to the Scene Task and eight to the Misbehaviour Task; all 12 faces were shown again during the Trait Recognition Task. Each child was photographed forward-facing against a neutral-coloured background with an off-white cape covering their clothes. The photographs were cropped to 420 x 420 pixels. All of the faces were categorized as emotionally neutral in a previous study (LoBue & Thrasher, 2015); we selected the identities that had been rated the highest (three male; three female) and lowest (three male; three female) on trustworthiness by a sample of adult participants (Collova et al., 2019)

Scene Task

Stimuli. Four CAFE identities were used in this task (two male: one high-trust, one low-trust; two female: one high-trust, one low-trust). Each face was superimposed into each of four ambiguous scenes (see Figure 1). Scenes were 720 x 480 pixels, and the face of the target child was approximately 112 x 112 pixels. Pilot testing ($n=25$) confirmed that each scene was ambiguous (i.e., the scenes were perceived no more positively than negatively) in the absence of the target child's face, single sample t-tests $ps > .07$. See *Coding Procedure* for more information on how the interpretations were coded for each scene.

Procedure. This task was adapted from a task previously used to investigate racial biases in children (McGlothlin & Killen, 2006; McGlothlin et al., 2005). In the present study, participants were presented with four hand-drawn scenes depicting an ambiguous situation between two children. In each trial, participants were first shown a photograph of a child's face for 3000ms with the child's name appearing above the image (e.g., This is Sally). The child's

face was then replaced with an ambiguous scene. Participants viewed two scenes in which the target child had a high-trust face (one female) and two scenes in which the target child had a low-trust face (one female). Four sets of scenes were created such that each scene was paired with each of the four faces; each participant was randomly assigned to one set and the order in which scenes were presented was randomized for each participant. The face of the second child was not visible (e.g., only the back of their head could be seen).

Participants provided three responses to each scene. 1) *Interpretation*. Participants were instructed to tell a story about what they believed to be happening in each situation and to base their stories on their instinctive response. 2) *Rating*. Participants were asked to rate the target child's behaviour using a 5-point child-friendly star scale ranging from 1 (*bad*) to 5 (*good*).

Participants were given the following child-friendly instructions: "How is [*child's name*] behaving in your story? Let's use our star scale! Please click anywhere between the smallest (1) and largest (5) star". 3) *Friendship evaluation*. Participants were asked: "Would [*child's name*] be a good friend?" Participants selected "Yes", "No", or "I don't know".

Before starting the task, participants completed two practice trials designed to verify that the participants understood the task and were able to use the response scale appropriately. The child faces used in the practice trials were hand drawn to depict one child smiling and one child frowning. All participants rated the behaviour of the target child in the positive scene (target child was smiling) more positively than the behaviour of the target child in the negative scene (target child was frowning).

Coding procedure. Two independent raters, blind to condition, coded the interpretations of each scene as positive (i.e., prosocial), neutral (i.e., ambiguous) and negative (i.e., antisocial). There was almost perfect agreement between the two raters, $\kappa = .83$, 95% CI [0.78, 0.87], p

< .001; disagreements were resolved by discussion. Examples of interpretations provided for Scene 1 (see Figure 1A) by young adult participants are as follows: *Matt is helping his friend build a castle of blocks. They are playing nicely.* (Positive); *Matt is playing with his friend, but he has a plan to destroy the tower his friend has been working on.* (Negative); *Matt and his friend are building a tower. He has a choice of whether to help add blocks or break the tower down.* (Neutral).

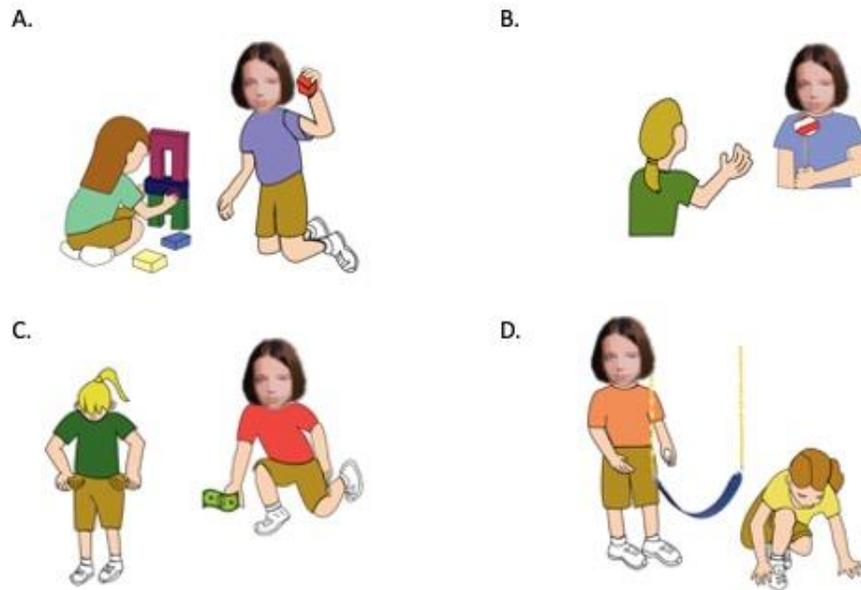


Figure 1. An example of the female version of four situations used in the Scene Task adapted from McGlothlin et al., 2005. Each scene depicts an ambiguous situation between two children. For example, in Scene A, is the target child helping to build the tower or about to knock the tower down? For copyright reasons, this is not an actual image of a child's face used in our experiment. The child face used in these examples was taken from the internet and their face has been blurred.

Misbehaviour Task

Stimuli. Eight novel CAFE identities were used in this task (four male: two high trust, two low trust; four female: two high trust, two low trust). The photographs were cropped to 420 x 420 pixels.

Procedure. Participants were shown a photograph of a child's face with their name appearing above the image for 3000ms. They then read a vignette about a misbehaviour (e.g., *Today at school, Sally spilled paint on their classmate's drawing*; see Table 1 for a complete list of the vignettes used in this task). The vignettes were worded such that the child's intentions were unclear (i.e., the behaviour might have been done on purpose or by accident). Participants were presented with four vignettes in which the target child had a high-trust face (two female) and four vignettes in which the target child had a low-trust face (two female). Four sets of vignettes were created such that each vignette was paired with two high- and two low-trust faces; each participant was randomly assigned to one set and the order in which vignettes were presented was randomized for each participant. Participants were asked to indicate whether the behaviour occurred on purpose or by accident and then completed the *Rating and Friendship Evaluation* questions from the Scene Task.

Before starting the task, participants completed two practice trials designed to verify that the participants understood the task and were able to use the response scale appropriately. The child faces used in the practice trials were hand drawn to depict one child smiling and one child frowning. All participants rated the behaviour of the target child in the positive vignette (target child was smiling) more positively than the behaviour of the target child in the negative vignette (target child was frowning).

Table 1. Vignettes presented in the Misbehaviour Task

Vignette	Test Question
Today during playtime, [<i>Child's Name</i>] broke their classmate's toy.	Did [<i>Child's Name</i>] break their classmate's toy on purpose or by accident?
Today during playtime, [<i>Child's Name</i>] spilled paint on their classmate's drawing.	Did [<i>Child's Name</i>] spill paint on their classmate's drawing on purpose or by accident?
Today in the classroom, [<i>Child's Name</i>] stepped on their classmate's foot.	Did [<i>Child's Name</i>] step on their classmate's foot on purpose or by accident?
Today in the classroom, [<i>Child's Name</i>] dropped their classmate's art project.	Did [<i>Child's Name</i>] drop their classmate's art project on purpose or by accident?
Today on the playground, [<i>Child's Name</i>] kicked a ball that hit their classmate on their head.	Did [<i>Child's Name</i>] hit their classmate with the ball on purpose or by accident?
Today on the playground, [<i>Child's Name</i>] skipped their classmate's turn in a game.	Did [<i>Child's Name</i>] skip their classmate's turn in the game on purpose or by accident?
Today at school, [<i>Child's Name</i>] knocked over the card tower that their classmate was building.	Did [<i>Child's Name</i>] knock over their classmate's card tower on purpose or by accident?
Today at school, [<i>Child's Name</i>] took their classmate's place in line.	Did [<i>Child's Name</i>] take their classmate's place in line on purpose or by accident?

Trait Recognition Task

Stimuli. The twelve CAFE faces used in the Scene and Misbehaviour Tasks (six high trust, three female; six low trust, three female) served as stimuli.

Procedure. Participants were asked to indicate how “nice” each of the 12 children appeared using a 5-point star scale ranging from 1 (*Mean*) to 5 (*Nice*); the order of presentation was randomized for each participant. We used the term “nice” as opposed to “trustworthy” in anticipation of testing children with this protocol. To verify that participants were paying attention, participants were asked to click the star that means “*Nice*” and the star that means “*Mean*”. Participants ($n=2$) who responded incorrectly to both questions were excluded.

Results

Analytic approach

To obtain ratings of trustworthiness in the Trait Recognition Task, the star scale was converted to a continuous scale ranging from 0-100, where 50 represented the midpoint of the scale (center of star 3: neutral). The dependent variable for the Trait Recognition Task was the location along the horizontal plane where participants clicked on the star scale. To examine whether attributions made about the ambiguous situations in the Scene and Misbehaviour Task reflected the facial trustworthiness of the target child, we conducted a mixed 2 (Trustworthiness: high, low) x 2 (Age: young adult, older adult) ANOVA on each dependent measure:

Interpretation, Behavioural Rating, Friendship Evaluation. The dependent variable for *Interpretation* was the proportion of trials in which participants provided a positive interpretation (Scene Task: providing a prosocial description; Misbehaviour Task: indicating that the behaviour was done by accident). For the *Behavioural Rating* measure, the star scale was converted to a continuous scale ranging from 0-100, where 50 represented the midpoint of the scale (center of

star 3: neutral). The dependent variable for the *Behavioural Rating* measure was the location along the horizontal plane where the participant clicked on the star scale. The dependent variable for the *Friendship Evaluation* measure was the proportion of trials in which participants indicated that the target child would be a good friend.

Trait Recognition Task

A 2 (Trustworthiness: high, low) x 2 (Age: young adult, older adult) analysis of variance (ANOVA) revealed a significant main effect of trustworthiness, $F(1, 116) = 1015.05, p < .001, \eta_p^2 = 0.49$. The target identities previously rated as high in trustworthiness (see Collova et al., 2019) were rated as more trustworthy ($M = 0.73, SE = 0.01$) than the target identities previously rated as low in trustworthiness ($M = 0.32, SE = 0.09$). There was also a significant main effect of age, $F(1, 116) = 9.47, p = .003, \eta_p^2 = .076$; older adults rated the identities more positively ($M = 0.54, SE = 0.01$) than young adults ($M = 0.51, SE = 0.01$). The interaction was not significant, $p = .278$, indicating that the effect of facial trustworthiness did not differ significantly across age groups. Single-sample t-tests (collapsed across age groups) confirmed that participants' ratings of high- and low-trust targets differed from the midpoint of the scale (50) in the expected directions, $ps < .001$, validating our stimuli (see Table 2).

Scene Task

Interpretation. Participants rarely provided a neutral or ambiguous interpretation (< 11% of all trials). As shown in Figure 2, participants were biased to provide an interpretation that matched the facial trustworthiness of the target child. Our analysis revealed a main effect of trustworthiness, $F(1, 116) = 110.73, p < .001, \eta_p^2 = .49$. Participants were more likely to provide a positive interpretation for situations in which the target child had a high-trust face ($M = 0.69, SE = 0.03$) as compared to situations in which the target child had a low-trust face ($M = 0.30, SE$

= .03). Neither the main effect of age nor the interaction was significant, $ps > .300$. Single-sample t-tests confirmed that participants provided a positive interpretation at above chance levels when the target child had a high-trust face and below chance levels when the target child had a low-trust face, $ps < .001$ (see Table 2).

Behavioural Rating. Our analysis revealed a main effect of trustworthiness, $F(1, 116) = 55.28, p < .001, \eta_p^2 = 0.32$. Participants rated the behaviour of high-trust targets more positively ($M = 0.59, SE = 0.16$) than low-trust targets ($M = 0.42, SE = 0.16$). Neither the main effect of age nor the interaction was significant, $ps > .265$. Single-sample t-tests confirmed that participants' ratings of high- and low-trust targets differed from the midpoint of the scale (50) in the expected directions, $ps < .001$ (see Table 2).

Friendship Evaluation. Our analysis revealed a main effect of trustworthiness, $F(1, 116) = 47.23, p < .001, \eta_p^2 = 0.29$. Participants were more likely to indicate that the target child would be a good friend when the target child had a high- ($M = 0.49, SE = 0.04$) as compared to a low-trust face ($M = 0.20, SE = 0.03$). Neither the main effect of age nor the interaction was significant, $ps > .062$. Single-sample t-tests revealed that participants provided a positive friendship evaluation below chance levels when the target child had a low-trust face, $p < .001$; positive friendship evaluations did not differ from chance levels when the target child had a high-trust face, $p = .630$ (see Table 2).

Misbehaviour Task

Interpretation. As shown in Figure 3, participants were biased to provide an interpretation that matched the facial trustworthiness of the target child. Our analysis revealed a main effect of trustworthiness, $F(1, 116) = 50.95, p < .001, \eta_p^2 = .31$. Participants were more likely to provide a positive interpretation (i.e., to say that the behaviour occurred by accident) for

situations in which the target child had a high-trust face ($M = 0.71$, $SE = 0.03$) as compared to situations in which the target child had a low-trust face ($M = 0.42$, $SE = 0.03$). Neither the main effect of age nor the interaction was significant, $ps > .168$. Single-sample t-tests revealed that participants provided positive interpretations above chance levels when the target child had a high-trust face, $p < .001$, and below chance levels when the target child had a low-trust face, $p = .003$ (see Table 2).

Behavioural Rating. Our analysis revealed a main effect of trustworthiness, $F(1, 116) = 55.22$, $p < .001$, $\eta_p^2 = 0.31$. Participants rated high-trust targets more positively ($M = 0.43$, $SE = 0.01$), than low-trust targets ($M = 0.32$, $SE = 0.01$). Neither the main effect of age nor the interaction was significant, $ps > .137$. Single-sample t-tests revealed that ratings of high- and low-trust targets were below the midpoint of the scale (50), $ps < .001$ (see Table 2).

Friendship Evaluation. Our analysis revealed a main effect of trustworthiness, $F(1, 116) = 46.62$, $p < .001$, $\eta_p^2 = 0.29$, and age group, $F(1, 116) = 5.49$, $p = .021$, $\eta_p^2 = 0.05$. Participants were more likely to indicate that the target child would be a good friend when the target child had a high- ($M = 0.35$, $SE = 0.03$) as compared to a low-trust face ($M = 0.13$, $SE = 0.02$). Older adults ($M = 0.29$, $SE = 0.03$) were more likely than young adults ($M = 0.18$, $SE = 0.03$) to indicate that the target child would be a good friend. There was no significant interaction, $p = .076$, indicating that the effect of facial trustworthiness on friendship evaluations did not differ significantly for young and older adults. Single-sample t-tests revealed that participants indicated that the target child would be a good friend below chance levels when the target child had a high- and low-trust face, $ps < .001$ (see Table 2).

Table 2. Results of single samples t-test from Experiment 1 and 2

Experiment 1: Young and Older Adults (<i>N</i> = 118)												
	High Trust						Low Trust					
	<i>M</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>LCI</i>	<i>UCI</i>	<i>M</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>LCI</i>	<i>UCI</i>
Trait Recognition Task	0.73	0.01	26.07	.000	0.71	0.75	0.32	0.01	-20.33	.000	0.30	0.34
Scene Task												
Interpretation	0.69	0.03	6.49	.000	0.63	0.75	0.30	0.03	-7.17	.000	0.24	0.36
Behavioural Rating	0.59	0.02	5.44	.000	0.55	0.63	0.42	0.02	-5.35	.000	0.38	0.46
Friendship Evaluation	0.48	0.04	-0.48	.630	0.40	0.56	0.20	0.03	-10.23	.000	0.14	0.26
Misbehaviour Task												
Interpretation	0.71	0.03	7.96	.000	0.65	0.77	0.42	0.03	-2.30	.003	0.36	0.48
Behavioural Rating	0.43	0.01	-5.91	.000	0.41	0.45	0.32	0.01	-17.77	.000	0.30	0.34
Friendship Evaluation	0.35	0.03	-4.51	.000	0.29	0.41	0.13	0.02	-18.63	.000	0.09	0.17
Experiment 2: Young Adults (<i>N</i> = 59)												
	High Trust						Low Trust					
	<i>M</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>LCI</i>	<i>UCI</i>	<i>M</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>LCI</i>	<i>UCI</i>
Trait Recognition Task	0.59	0.01	7.19	.000	0.57	0.60	0.42	0.01	-7.10	.000	0.40	0.44
Scene Task												
Interpretation	0.59	0.04	2.10	.040	0.51	0.67	0.40	0.04	-2.35	.022	0.32	0.48
Behavioural Rating	0.56	0.02	2.74	.008	0.52	0.60	0.49	0.02	-0.35	.729	0.45	0.53
Friendship Evaluation	0.48	0.05	-0.55	.582	0.38	0.58	0.38	0.04	-2.79	.007	0.30	0.46
Misbehaviour Task												
Interpretation	0.66	0.04	4.46	.000	0.58	0.74	0.48	0.04	-0.41	.684	0.40	0.56
Behavioural Rating	0.39	0.01	-8.57	.000	0.37	0.41	0.32	0.01	-12.59	.000	0.30	0.34
Friendship Evaluation	0.26	0.04	-5.37	.000	0.18	0.34	0.16	0.03	-12.62	.000	0.10	0.22

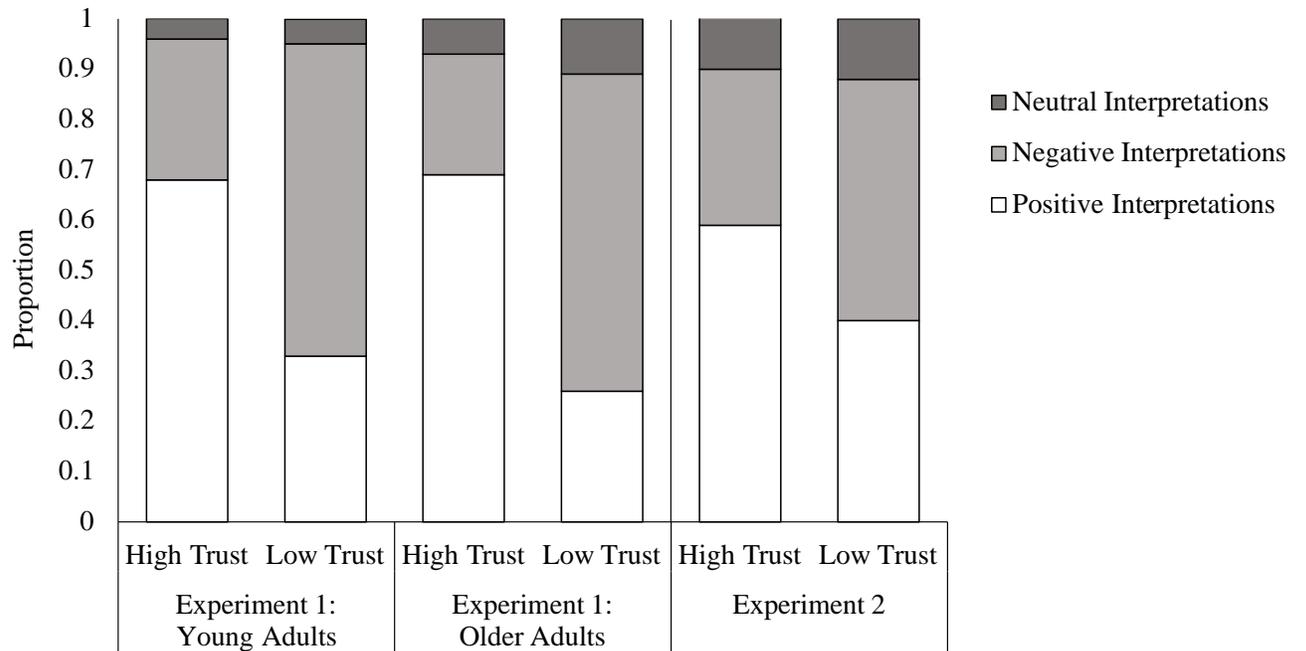


Figure 2. Proportion of participants in Experiment 1 and 2 who did (white bars) and did not (sum of grey bars) provide a positive interpretation in the Scene Task for targets with a high- vs. low-trust face.

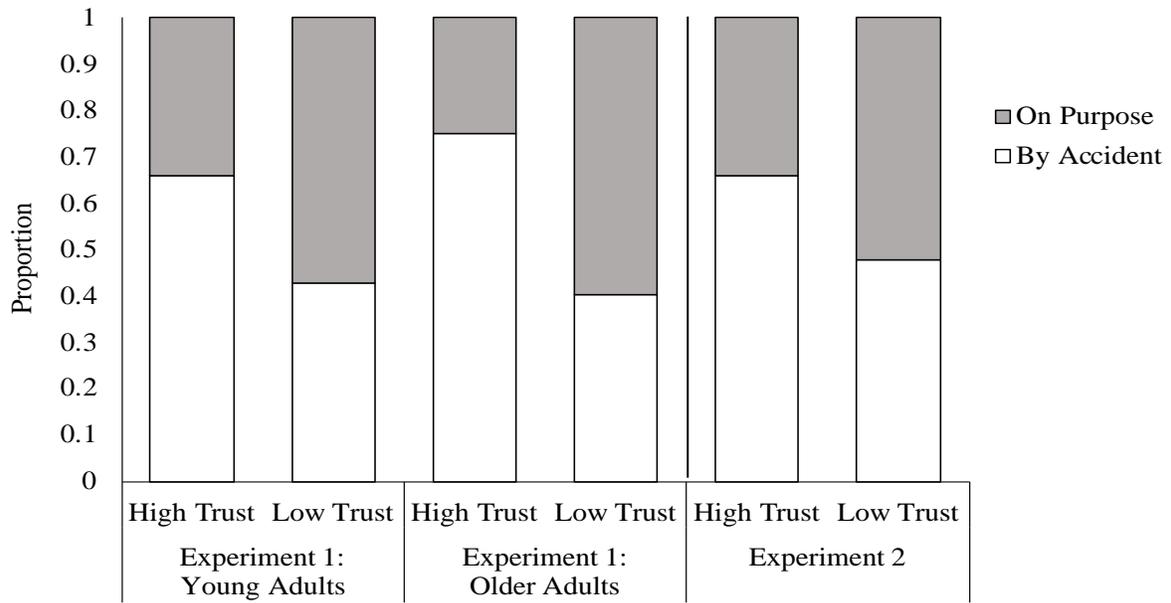


Figure 3. Proportion of participants in Experiment 1 and 2 who did (white bars) and did not (grey bars) provide a positive interpretation in the Misbehaviour Task (i.e., indicate that the behaviour happened by accident, as opposed to on purpose) for targets with a high- vs. low-trust face.

Discussion

Our results demonstrate that facial trustworthiness influenced young and older adults' interpretations of children's behaviour in ambiguous situations. Ambiguous scenes were more likely to be interpreted positively when the target child had a high-trust face, as opposed to a low-trust face (i.e., the situation was more likely to be given a prosocial explanation). Likewise, both young and older adults were more likely to infer that misbehaviours were accidental if the target child had a high-trust face than a low-trust face. The behaviour of the target child in both tasks was rated more positively when the target child had a high- as opposed to a low-trust face and participants were more inclined to indicate that the target child would be a good friend when the target child had a high- as opposed to a low-trust face.

Young and older adults showed comparable patterns with only two exceptions. Overall, older adults rated child faces higher in trustworthiness in the Trait Recognition Task and were more likely to indicate that the target child would be a good friend in the Misbehaviour Task as compared to young adults. Nonetheless, in no case did the effect of facial trustworthiness vary with participant age. These results suggest that trustworthiness in child faces likely has real-world consequences. Children with high-trust faces are more likely to receive the benefit of the doubt as compared to children with low-trust faces when the child's intent is not clear and their actions are uncertain (i.e., Scene Task) or a misbehaviour has been performed (i.e., Misbehaviour Task).

Our findings indicate that the Scene and Misbehaviour tasks capture different aspects of first impressions. In both tasks, facial trustworthiness influenced perceived intent; participants provided positive interpretations above chance levels when the target child had a high-trust face. However, in the Scene Task, where the behaviour was ambiguous, participants rated the

behaviour of high-trust targets above the midpoint of the scale (i.e., more positively), and friendship evaluations did not differ from chance. In contrast, in the Misbehaviour Task, where the intention behind the misbehaviour was ambiguous, participants rated the behaviour of high-trust targets below the midpoint of the scale (i.e., more negatively) and friendship evaluations of high-trust targets were below chance levels. These results suggest that adults' perceptions of children's behaviour and the likelihood of that child being a good friend depend not only on the child's intention, but also on whether the child's behaviour is ambiguous (Scene Task) or whether a misbehaviour has been performed (Misbehaviour Task).

Prior to drawing strong conclusions, we conducted a second experiment to replicate our findings using a more controlled set of stimuli to verify that the effects were driven by differences in facial trustworthiness. Given comparable performance in young and older adults in Experiment 1 we only tested young adults in Experiment 2.

Experiment 2

The main purpose of Study 2 was to replicate the findings of Study 1 while also controlling for the identity of the child. We used facial images that had been transformed so as to systematically vary in trustworthiness. As in Experiment 1, this set of stimuli was previously validated by a sample of adult participants (Collova et al., 2019). Whereas Collova et al. (2019) reported that these stimuli differed in trustworthiness and influenced adults' behavioural expectations of children (i.e., which of two children should be rewarded versus watched more closely for rough play), here we examined whether controlled differences in trustworthiness influence adults' interpretation of children's intention to help versus harm.

Method

Participants

Fifty-nine Caucasian young adults (age range: 18-27; $M = 23.42$, $SD = 2.47$; 33 female) were recruited online through Testable Minds (Testable S.R.L., 2018), a sample size that matches that of Experiment 1. G*Power 3.1 confirmed that this sample size exceeded that required to find a medium effect size ($t=1.71$) in a paired samples t-test with power = 0.8. Participants were compensated \$2.00 USD.

General Procedure

The entire battery of tasks was completed online on Qualtrics (Qualtrics, Provo, UT). The procedure was identical to that of Study 1 and received clearance from the Research Ethics Board at Brock University.

Face Stimuli

Two versions (high- vs. low-trust) of 12 Caucasian child faces from the CAFE Set (LoBue, 2014; LoBue & Thrasher, 2015) were obtained from Jemma Collova through personal communication. To create these versions, a high- and low-trust face average were created by morphing together the five child faces from the sample of CAFE faces that had been rated the highest or lowest in trustworthiness respectively. Then, the original images of the 12 faces that had average trustworthiness ratings were morphed with both the high- and low-trust face average, creating a 50% high- and 50% low-trust version for each identity (*see* Collova et al., 2019: Study 4 for more information on how the face averages were created).

The original images of these 12 faces had average trustworthiness ratings. Four identities were assigned to the Scene Task, and eight were assigned to the Misbehaviour Task. The situation within each task to which each identity was assigned and which version (high- vs. low-

trust) was shown was counterbalanced across participants; for each participant, the version shown in the Trait Recognition Task matched that shown in the Ambiguous Situations protocol. The stimuli were the same size as Experiment 1.

Results

Analytic Approach

To examine ratings of trustworthiness in the Trait Recognition Task, we conducted a paired samples t-test. To examine whether attributions made about the ambiguous situations in the Scene and Misbehaviour Task reflected the facial trustworthiness of the target child, we conducted paired samples t-tests on each dependent measure: *Interpretation*, *Behavioural Rating*, *Friendship Evaluation*. The dependent variables were the same as for Experiment 1.

Trait Recognition Task

A paired samples t-test revealed that target identities morphed towards the high-trust composite ($M = 0.59$, $SE = 0.01$) were rated as more trustworthy than the same target identities morphed towards the low-trust composite ($M = 0.42$, $SE = 0.01$), $t(58) = -10.84$, $p < .001$, $d = 1.41$. Single-sample t-tests confirmed that ratings of high- and low-trust faces differed from the midpoint of the scale (50) in the expected directions, $ps < .001$, validating our stimuli (see Table 2).

Scene Task

Interpretation. As shown in Figure 2, participants were biased to provide an interpretation that matched the facial trustworthiness of the target child. Participants were more likely to provide a positive interpretation for situations in which the target child was presented with a high-trust face ($M = 0.59$, $SE = 0.04$) as compared to situations in which the target child was presented with a low-trust face ($M = .40$, $SE = .04$), $t(58) = 3.55$, $p = .001$, $d = 0.46$. Single-

sample t-tests confirmed that participants provided a positive interpretation at above chance levels when the target child was presented with a high-trust face, $p = .040$, and below chance levels when the target child was presented with a low-trust face, $p = .022$ (see Table 2).

Behavioural Rating. Participants rated the behaviour of targets more positively when the target child was presented with a high- ($M = 0.56$, $SE = 0.02$) as compared to a low-trust face ($M = 0.49$, $SE = 0.02$), $t(58) = 2.46$, $p = .017$, $d = 0.32$. Single-sample t-tests revealed that targets presented with a high-trust face were rated above the midpoint of the scale (50), $p = .008$; rating of targets presented with a low-trust face did not differ from the midpoint of the scale (50), $p = .729$ (see Table 2).

Friendship Evaluation. Our analysis revealed no significant difference between participants' indications that the target child would be a good friend when the target child was presented with a high- ($M = 0.48$, $SE = 0.05$) as compared to a low-trust face ($M = 0.38$, $SE = 0.04$), $t(58) = 1.75$, $p = .086$, $d = 0.23$. Single-sample t-tests revealed that participants provided a positive friendship evaluation below chance levels when the target child was presented with a low-trust face, $p = .007$; positive friendship evaluations did not differ from chance levels when the target child was presented with a high-trust face, $p = .582$ (see Table 2).

Misbehaviour Task

Interpretation. As shown in Figure 3, participants were biased to provide an interpretation that matched the facial trustworthiness of the target child. Participants were more likely to provide a positive interpretation (i.e., claim that the misbehaviour was unintentional) for situations in which the target child was presented with a high- ($M = 0.66$, $SE = 0.04$) as compared to a low-trust face ($M = 0.48$, $SE = 0.04$), $t(58) = 3.57$, $p = .001$, $d = 0.46$. Single-sample t-tests revealed that participants provided positive interpretations at above chance levels

when the target child was presented with a high-trust face, $p < .001$; positive interpretations did not differ from chance levels when the target child was presented with a low-trust face, $p = .684$ (see Table 2).

Behavioural Rating. Participants rated targets presented with high-trust faces more positively ($M = 0.39$, $SE = 0.01$) than low-trust targets ($M = 0.32$, $SE = 0.01$), $t(58) = 4.87$, $p < .001$, $d = 0.63$. Single-sample t-tests revealed that participants' ratings of high- and low-trust faces were below the midpoint of the scale (50), $ps < .001$ (see Table 2).

Friendship Evaluation. Participants were more likely to indicate that the target child would be a good friend when the target child was presented with a high- ($M = 0.26$, $SE = 0.04$) as compared to a low-trust face ($M = 0.16$, $SE = 0.03$), $t(58) = 2.43$, $p = .018$, $d = 0.32$. Single-sample t-tests revealed that participants indicated that the target child would be a good friend at below chance levels for both high- and low-trust targets, $p < .001$ (see Table 2).

Discussion

Our results from Experiment 2 replicate the overall pattern from Experiment 1 and confirm that adults are sensitive to subtle differences in trustworthiness in child faces and that these impressions of trustworthiness influence adults' interpretations of children's behaviour in ambiguous situations. As in Experiment 1, high-trust targets were more likely to receive a pro-social interpretation than low-trust targets in the Scene Task and their behaviour was rated more favourably. As in Experiment 1, misbehaviours of high-trust targets were more likely to be treated as accidental, the behaviour of high-trust targets was rated more positively, and participants were more inclined to indicate that high-trust targets would be a good friend. Thus, even when controlling for facial identity, adults' interpretations of ambiguous situations were

influenced by the facial trustworthiness of the target child. Collectively, these experiments provide strong evidence that facial trustworthiness has real-world consequences for children.

General Discussion

Adults' perceptions of children can have profound implications for daily social interactions, just as they do for adults (Bagnis et al., 2020; Ballew & Todorov, 2007; Castelli et al., 2009; Centorrino et al., 2015; Jaeger et al., 2020; Mattarozzi et al., 2017; Olivola & Todorov, 2010; Rule & Ambady, 2008; 2011; Todorov et al., 2005; Wilson & Rule, 2015). Previous research has found that adults' perceptions of child faces influence adults' expectations of children such that when shown a pair of child faces (one high- and one low-nice), adults are more likely to select the child with the high-nice face to be rewarded for good behaviour as compared to the child with a low-nice face (Collova et al., 2019). These first impressions of child faces are likely to have cascading effects on development; children perceived as "nice" are likely to enjoy more positive interactions with adults and come to behave more positively as a result (Merton, 1948). The present study examined adults' perceptions of children's behaviour in ambiguous situations—a context in which first impressions have never been examined for adult or child faces despite closely reflecting adults' use of first impressions in daily life.

Our novel Ambiguous Situations protocol generated three main findings. Our first major finding supported our hypothesis that young and older adults would be biased to provide an interpretation for the ambiguous situations that matched the facial trustworthiness of the target child. Ambiguous scenes were more likely to be interpreted as prosocial when the target child had a high-trust as compared to a low-trust face; misbehaviours were more likely to be interpreted as unintentional when the target child had a high-trust as compared to a low-trust face. Furthermore, the child's behaviour was rated as more positive when the child had a high-

trust-face and these facial biases extended beyond immediate behavioural attributions to stable characteristics of the child: children with high-trust faces were more likely to be perceived as being a good friend as compared to children with low-trust faces.

Exploring adults' interpretations of ambiguous situations using two novel tasks provided unique advantages. The Scene Task reflects real-life use of first impressions. For example, on the playground, teachers often must evaluate the behaviour at hand based only on a quick glimpse of children's actions. In our task, participants were not explicitly cued to examine the trustworthiness of the target child's face; trustworthiness ratings were collected after the Ambiguous Situations protocol was complete. As a result, our task presents an ecologically valid assessment of adults' implicit trust biases of children's faces. The Misbehaviour Task presents another powerful assessment of adults' interpretations of ambiguous behaviour. In this task, the child did do something wrong and only the intention behind the act could vary. Here, our assessment sheds light on how children's misbehaviours are interpreted as a function of facial trustworthiness. Our results suggest that children with high-trust faces are more likely to be given the benefit of the doubt—a bias that is likely to result in differential treatment for children with high- versus low-trust faces.

Our second major finding was that controlled differences in trustworthiness (i.e., facial stimuli morphed towards a high- and low-trust face average) influenced adults' interpretations of children's behaviour. Whereas different identities represented high- and low-trust faces in Experiment 1, average-trust faces were manipulated to create high- and low-trust versions of the same identity in Experiment 2. Our results from Experiment 2 confirm that young adults are sensitive to subtle differences in trustworthiness, differences that are carried by subtle displays of emotion, babyfacedness, masculinity, weight, facial width to height ratio, perceived health, and

attractiveness (*see* Collova et al., 2019: Study 3). Using controlled stimuli provides a critical test of our claim that facial trustworthiness influences adults' interpretations of ambiguous situations—findings that support the robustness of our protocol. Future research could expand our current protocol to use more diverse facial stimuli. Our study only included photographs of Caucasian children and our sample of adults was primarily Caucasian. Doing so eliminated face-race biases from influencing our findings; previous research has shown that culture-specific face typicality influences perceptions of trustworthiness (Sofer et al., 2017). Future research should examine to what extent our first impressions of child faces are universal versus specific to own- versus other-race faces by using a more diverse sample of participants and facial stimuli. Further, the child faces in our study were between the ages of 4-6; now that we have established this protocol, future studies could investigate interpretations of ambiguous behaviour across age groups. In addition, our tasks could extend to other situations including lie- versus truth-telling and cheating versus honest performance on a test to examine adults' perceptions of children's behaviour across a wide range of contexts.

Our third major finding was that the influence of facial trustworthiness was comparable for young and older adults. In support of previous research demonstrating that older adults are more positive than young adults when rating trustworthiness of adult faces (Zebrowitz et al., 2013), we found that older adults were more positive in their ratings of child faces, although the effect of facial trustworthiness was comparable for the two age groups. Counter to predictions, older adults' performance in the Scene and Misbehaviour Task was comparable to that of young adults. The only significant difference between the two age groups was that older adults were more inclined than young adults to indicate that the target child would be a good friend in the Misbehaviour Task; whereas young adults' ratings were below chance levels, older adults'

ratings did not differ from chance levels on this variable. This suggests that when presented with a child's misbehaviour, older adults were less inclined to use this situational information to bias their stable trait inferences (i.e., in deciding whether the child is a good friend). It is interesting to consider what mechanisms are underlying age-related differences in adults' perceptions of children's facial trustworthiness. The Socioemotional Selectivity Theory posits that with age, there is a motivational shift associated with an increased preference for positive as opposed to negative stimuli (Carstensen et al., 1999; Carstensen & Mikels, 2005). Contrary to past research showing that older adults provide more positive evaluations of negative stimuli than young adults (Carstensen & Mikels, 2005; Mather & Carstensen, 2005; O'Connor et al., 2019; Zebrowitz et al., 2013), we found no evidence that older adults provided more positive ratings or interpretations of behaviour for low-, as opposed to high-trust faces. We do not know if our lack of effect is attributable to the use of child faces (i.e., perhaps the effect would be stronger when using adult faces—faces that would appear more threatening to older adults), low cognitive demands (Zebrowitz et al., 2017), or alternatively, if the positivity bias does not extend to ambiguous behaviour. Nonetheless, we provide convincing evidence that children's facial trustworthiness matters for adults of all ages.

Functional accounts suggest that first impressions of trustworthiness are an evolutionary adaptation to detect others' intention to help versus harm (Oosterhof & Todorov, 2008). Building directly upon the functional account, our protocol provides insight on the influence of facial trustworthiness on adults' interpretations of ambiguous situations—situations that are especially sensitive to intent. Our ambiguous scenes are sensitive to how facial trustworthiness influences adults' interpretation of ongoing behaviour; our ambiguous vignettes are sensitive to how facial trustworthiness influences adults' interpretation of intent behind misbehaviour. This is an ideal

setting in which to examine perceptions of someone's intention to help versus harm because it is precisely the intentions of the perpetrator that are unclear, aligning our protocol closely with functional accounts of first impressions (Oosterhof & Todorov, 2008; Collova et al., 2019). Children do not pose threat to adults (Collova et al., 2019) and so the effects observed in the current experiments reflect their perception of niceness (i.e., the child's intent to help vs. harm someone else).

Our novel protocol also provides a unique opportunity to examine the development of children's sensitivity to facial trustworthiness. We are currently testing child participants on this child-friendly protocol to examine how children's sensitivity to facial trustworthiness influences interpretations of peer behaviour in ambiguous situations. Although children demonstrate (near) adult-like consistency in attributing trait labels such as "nice" and "mean" to both adult and child faces that vary in trustworthiness (Cogsdill & Banaji, 2015; Cogsdill et al., 2014), there are mixed findings as to whether children's performance is adult-like when using these facial cues to decide how to behave towards others (Antonakis & Dalgas, 2009; Bascandziev & Harris, 2016; Charlesworth et al., 2019; Clément et al., 2013; Ewing et al., 2015; Ewing et al., 2019; Mondloch et al., 2019; Palmquist et al., 2019; Palmquist & DeAngelis, 2020; Rosati et al., 2019). Most research has examined children's first impressions of adult, rather than child faces. In line with the functional account, a peer setting might better capture children's use of first impressions in daily life as much of a child's social learning occurs at school (e.g., when detecting threat among peers). If perceptions of trustworthiness are adaptive, children should be sensitive to trustworthiness in ambiguous situations where it is particularly functional to detect help versus harm (e.g., when identifying a friend versus bully on the playground).

In anticipation of testing children on this protocol, in the Scene and Misbehaviour task, participants completed two practice trials. The child faces used in the practice trials were hand drawn to depict one child smiling and one child frowning. These overt expressions were designed to verify that participants understood the task. However, it is possible that the cartoon faces inadvertently primed participants' attention to the facial cues. We are currently replicating our protocol using additional child faces that are neutral in expression and previously rated as average in trustworthiness as opposed to cartoon faces in the practice trials to explore whether exposure to overt emotional expressions influences performance on our task.

Finally, facial appearance has been shown to play a fundamental role in influencing adults' perceptions of children. It would be interesting to expand our novel protocol to investigate how young and older adults' perceptions of children's behaviour influences their behaviour *towards* children (i.e., interfering with children's ongoing behaviour, deciding severity of punishment). Adult and child judgements of children's trustworthiness were predictive of children's real-world trustworthiness as reported by their classmates (Li et al., 2017). Children who were perceived as more trustworthy as judged by unfamiliar adult and child participants were also rated as more trustworthy by classmates and rated higher in peer acceptance. For the most part, in adult literature first impressions are not accurate (Olivola & Todorov, 2010b). The findings from Li and colleagues (2017) suggest that there might be a hint of truth to perceptions of trustworthiness in child faces, perhaps a result of the self-fulfilling prophecy unfolding (i.e., children with high-trust faces being treated differently and in turn, fulfilling others' expectations). Given the real-life association, it is critical to understand how facial trustworthiness influences how children are treated by adults and their peers.

In conclusion, this is the first study to examine the influence of facial trustworthiness on adults' interpretation of child behavior in an ambiguous context. Our results provide strong evidence that facial appearance biases adults' perceptions of children. When presented with a child's face high in trustworthiness, young and older adults were more likely to interpret children's ambiguous behaviour as pro-social, rate the child's behaviour more positively, and indicate that the child would be a good friend. These results demonstrate the influence of facial trustworthiness of child faces at biasing perceptions of children's behaviour—findings that are at risk of posing real-life behavioural consequences for children.

References

- Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play! *Science*, *323*, 1183.
<https://doi.org/10.1126/science.1167748>
- Aradhye, C., Vonk, J., & Arida, D. (2015). Adults' responsiveness to children's facial expressions. *Journal of Experimental Child Psychology*, *135*, 56–71.
<http://dx.doi.org/10.1016/j.jecp.2015.02.006>
- Bagnis, A., Caffo, E., Cipolli, C., De Palma, A., Farina, G., & Mattarozzi, K. (2020). Judging health care priority in emergency situations: Patient facial appearance matters. *Social Science & Medicine*, *260*, 113-180. <https://doi.org/10.1016/j.socscimed.2020.113180>
- 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>
- Bascandziev, I., & Harris, P. L. (2016). The beautiful and the accurate: Are children's selective trust decisions biased? *Journal of Experimental Child Psychology*, *152*, 92–105.
<https://doi.org/10.1016/j.jecp.2016.06.017>
- Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist*, *54*(3), 165-181.
<http://doi.org/10.1037/0003-066X.54.3.165>
- Carstensen, L. L., & Mikels, J. A. (2005). At the intersection of emotion and cognition: Aging and the positivity effect. *Current Directions in Psychological Science*, *14*(3), 117–121. <http://doi.org/10.1111/j.0963-7214.2005.00348.x>

- Castelli, L., Carraro, L., Ghitti, C., & Pastore, M. (2009). The effects of perceived competence and sociability on electoral outcomes. *Journal of Experimental Social Psychology, 45*(5), 1152–1155. <https://doi.org/10.1016/j.jesp.2009.06.018>
- 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>
- Centorrino, S., Djemai, E., Hopfensitz, A., Milinski, M., & Seabright, P. (2015). Honest signaling in trust interactions: Smiles rated as genuine induce trust and signal higher earning opportunities. *Evolution and Human Behavior, 36*(1), 8–16. <https://doi.org/10.1016/j.evolhumbehav.2014.08.001>
- Charlesworth, T. E. S., Hudson, S., Kiera T. J., 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), 1400–1413. <https://doi.org/10.1037/dev0000734>
- Clément, F., Bernard, S., Grandjean, D., & Sander, D. (2013). Emotional expression and vocabulary learning in adults and children. *Cognition and Emotion, 27*(3), 539–548. <https://doi.org/10.1080/02699931.2012.724012>
- Clifford, M. M., & Walster, E. (1973). The effect of physical attractiveness on teacher expectations. *Sociology of Education, 46*, 248–258. <http://dx.doi.org/10.2307/2112099>
- 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>
- Collova, J. R., Sutherland, C. A., & 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, 117*(5), 900-924.
<https://doi.org/10.1037/pspa0000167>
- Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2015). Perceived trustworthiness of faces drives trust behaviour in children. *Developmental Science, 18*(2), 327–334.
<https://doi.org/10.1111/desc.12218>
- Ewing, L., Sutherland, C. A. M., & Willis, M. L. (2019). Children show adult-like facial appearance biases when trusting others. *Developmental Psychology, 55*(8), 1–8.
<https://doi.org/10.1037/dev0000747>
- Harris, M. J., & Garris, C. P. (2008). *You never get a second chance to make a first impression: Behavioral consequences of first impressions*. In N. Ambady & J. J. Skowronski (Eds.), *First impressions* (p. 147–168). Guilford Publications.
- Jaeger, B., Todorov, A. T., Evans, A. M., & van Beest, I. (2020) 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>
- Kenealy, P., Frude, N., & Shaw, W. (1988). Influence of children's physical attractiveness on teacher expectations. *The Journal of Social Psychology, 128*(3), 373-383.
<http://dx.doi.org/10.1080/00224545.1988.9713754>

- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*, 390–423. <http://dx.doi.org/10.1037/0033-2909.126.3.390>
- Li, Q., Heyman, G. D., Mei, J., & Lee, K. (2019). Judging a book by its cover: Children’s facial trustworthiness as judged by strangers predicts their real-world trustworthiness and peer relationships. *Child Development*, *90*, 562–575. <http://dx.doi.org/10.1111/cdev.12907>
- LoBue, V. (2014). *The Child Affective Facial Expression (CAFE) set*. Databrary. <http://doi.org/10.17910/B7301K>
- LoBue, V., & Thrasher, C. (2015). The Child Affective Facial Expression (CAFE) set: validity and reliability from untrained adults. *Frontiers in Psychology*, *5*, 1532. <http://doi.org/10.3389/fpsyg.2014.01532>
- Mather, M., & Carstensen, L. L. (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in cognitive sciences*, *9*(10), 496-502. <http://doi.org/10.1016/j.tics.2005.08.005>
- Mattarozzi, K., Colonnello, V., Gioia, F. De, & Todorov, A. (2017). I care, even after the first impression: Facial appearance-based evaluations in healthcare context. *Social Science & Medicine*, *182*, 68-72. <http://doi.org/10.1016/j.socscimed.2017.04.011>
- McGlothlin, H., & Killen, M. (2006) Intergroup attitudes of European American children attending ethnically homogeneous schools. *Child Development*, *77*, 1375–1386. <http://doi.org/10.1111/j.1467-8624.2006.00941.x>
- McGlothlin, H., Killen, M., & Edmonds, C. (2005). European-American children’s intergroup attitudes about peer relationships. *British Journal of Developmental Psychology*, *23*, 227–249. <http://doi.org/10.1348/026151005x26101>

- Merton, R. K. (1948). The self-fulfilling prophecy. *The antioch review*, 8(2), 193-210.
<https://www.jstor.org/stable/4609267>
- 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>
- O'Connor, A. M., Lyon, T. D., & Evans, A. D. (2019). Younger and older adults' lie-detection and credibility judgments of children's coached reports. *Psychology, Crime & Law*, 25(9), 925-944. <https://doi.org/10.1080/1068316X.2019.1597092>
- Olivola, C. Y., & Todorov, A. (2010a). 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>
- Olivola, C. Y., & Todorov, A. (2010b). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, 46(2), 315-324. <https://doi.org/10.1016/j.jesp.2009.12.002>
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, 105(32), 11087–11092.
<https://doi.org/10.1073/pnas.0805664105>
- 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*, 38, 31-41. <http://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, <https://doi.org/10.1016/j.cogdev.2020.100948>

- Rosati, A. G., Benjamin, N., Pieloch, K., & Warneken, F. (2019). Economic trust in young children. *Proceedings of the Royal Society B: Biological Sciences*, 286(1907), 20190822. <https://doi.org/10.1098/rspb.2019.0822>
- Rule, N., & Ambady, N. (2010). First impressions of the face: Predicting success. *Social and Personality Psychology Compass*, 4(8), 506–516. <https://doi.org/10.1111/j.1751-9004.2010.00282.x>
- Rule, N. O., & Ambady, N. (2011). Judgments of power from college yearbook photos and later career success. *Social Psychological and Personality Science*, 2(2), 154–158. <https://doi.org/10.1177/1948550610385473>
- Rule, N. O., & Ambady, N. (2008). The face of success: Inferences from chief executive officers' appearance predict company profits: Short report. *Psychological Science*, 19(2), 109–111. <https://doi.org/10.1111/j.1467-9280.2008.02054.x>
- Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*, 9(2), 260. <https://doi.org/10.1037/a0014681>
- Sofer, C., Dotsch, R., Oikawa, M., Oikawa, H., Wigboldus, D. H., & 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>
- 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., & Young, A. W. (2016). Integrating social and facial models of person perception: Converging and diverging dimensions. *Cognition*, *157*, 257–267. <https://doi.org/10.1016/j.cognition.2016.09.006>
- Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton University Press.
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623–1626. <https://doi.org/10.1126/science.1110589>
- Willis, J., & Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*, 592–98. <https://doi.org/10.1111/j.1467-9280.2006.01750.x>
- Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, *26*(8), 1325–1331. <https://doi.org/10.1177/0956797615590992>
- Zebrowitz, L. A., Boshyan, J., Ward, N., Gutchess, A., & Hadjikhani, N. (2017). The older adult positivity effect in evaluations of trustworthiness: Emotion regulation or cognitive capacity? *PloS one*, *12*(1), e0169823. <https://doi.org/10.1371/journal.pone.0169823>
- 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., Franklin, R. G., Hillman, S., & Boc, H. (2013). Older and younger adults' first impressions from faces: Similar in agreement but different in positivity. *Psychology and Aging*, *28*(1), 202–212. <https://doi.org/10.1037/a0030927>

Zebrowitz, L. A., Kendall-Tackett, K., & Fafel, J. (1991). The influence of children's facial maturity on parental expectations and punishments. *Journal of Experimental Child Psychology*, 52, 221–238. [http://dx.doi.org/10.1016/0022-0965\(91\)90060-6](http://dx.doi.org/10.1016/0022-0965(91)90060-6)

Zebrowitz, L. A., & Montepare, J. M. (2015). *Faces and first impressions*. In M. Mikulincer, P. R. Shaver, E. Borgida, & J. A. Bargh (Eds.), *APA handbooks in psychology®. APA handbook of personality and social psychology, Vol. 1. Attitudes and social cognition* (p. 251–276). American Psychological Association. <https://doi.org/10.1037/14341-008>