British Journal of Psychology (2015)

The British Psychological Society

www.wileyonlinelibrary.com

Judging trustworthiness from faces: Emotion cues modulate trustworthiness judgments in young children

Frances Caulfield[1], Louise Ewing[1,2]*, Samantha Bank[1] and Gillian Rhodes[1]

[1]ARC Centre of Excellence in Cognition and its Disorders, School of Psychology, University of Western Australia, Crawley, Western Australia, Australia
[2]Department of Psychological Sciences, Birkbeck, University of London, UK

By adulthood, people judge trustworthiness from appearances rapidly and reliably. However, we know little about these judgments in children. This novel study investigates the developmental trajectory of explicit trust judgments from faces, and the contribution made by emotion cues across age groups. Five-, 7-, 10-year-olds, and adults rated the trustworthiness of trustworthy and untrustworthy faces with neutral expressions. The same participants also rated faces displaying overt happy and angry expressions, allowing us to investigate whether emotion cues modulate trustworthiness judgments similarly in children and adults. Results revealed that the ability to evaluate the trustworthiness of faces emerges in childhood, but may not be adult like until 10 years of age. Moreover, we show that emotion cues modulate trust judgments in young children, as well as adults. Anger cues diminished the appearance of trustworthiness for participants from 5 years of age and happy cues increased it, although this effect did not consistently emerge until later in childhood, that is, 10 years of age. These associations also extended to more subtle emotion cues present in neutral faces. Our results indicate that young children are sensitive to facial trustworthiness, and suggest that similar expression cues modulate these judgments in children and adults.

The belief that personality traits can be inferred from an individual's face can be traced back to the pseudo-science of physiognomy, in which human character is assessed from external physical features (Lavater, 1789). Today, people still infer personality traits from faces. For example, adults evaluate the trustworthiness of faces rapidly and with a high degree of consensus (Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006). Moreover, these judgments modulate behaviour, with trustworthy-looking individuals receiving larger 'investments' from partners in economic trust games than those that look untrustworthy (Chang, Doll, van't Wout, Frank, & Sanfey, 2010; Ewing, Caulfield, Read, & Rhodes, 2014; Rezlescu, Duchaine, Olivola, & Chater, 2012).

These appearance-based trust inferences may partially reflect transient facial cues, such as emotional expressions, being interpreted as signals of more enduring interpersonal attributes or intentions (McArthur & Baron, 1983; Secord, 1958; Zebrowitz & Collins, 1997). Researchers have proposed that trustworthiness judgments may be

*Correspondence should be addressed to Louise Ewing, School of Psychology, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia (email: louise.ewing@uwa.edu.au).

DOI:10.1111/bjop.12156

modulated by overt facial expressions, in a process known as temporal extension. Temporal extension describes the effect where 'the perceiver regards a momentary characteristic of a person as if it were an enduring attribute' (Secord, 1958). In addition, trustworthiness judgments may also be modulated by more subtle expressions, such as the resemblance of a person's permanent (resting) facial structure to emotional expressions. This process, whereby an individual is perceived to have those traits that are associated with the emotional expressions that their features resemble, is known as emotion overgeneralization (Zebrowitz, 1996, 1997).

A large body of evidence supports an association between perceived emotions and trustworthiness (Franklin & Zebrowitz, 2013; Oosterhof & Todorov, 2008, 2009; Said, Sebe, & Todorov, 2009; Sutherland *et al.*, 2013; Zebrowitz, Kikuchi, & Fellous, 2010). Faces displaying cues to happiness consistently appear relatively trustworthy and those displaying cues to anger appear relatively untrustworthy, with this effect persisting even when authors control for the effects of other factors that might also influence trait impressions, such as attractiveness and babyfacedness (Franklin & Zebrowitz, 2013; Zebrowitz *et al.*, 2010). Computer-modelling studies also show that manipulating the trustworthiness of faces influences their perceived emotional expressions. Exaggerating computer-generated faces to appear extremely untrustworthy increases the likelihood of these faces being classified as angry by adult participants and making faces appear extremely trustworthy increases their likelihood of being classified as happy, by adult participants (Oosterhof & Todorov, 2008; Sutherland *et al.*, 2013). These results strongly support a link between perceived trustworthiness and facial cues to emotion.

It is perhaps surprising, given the considerable research interest in facial trust perception, that few studies have investigated this trait impression outside of the typical, young adult population. Franklin and Zebrowitz (2013) identified intact overgeneralization of anger to judgments of trustworthiness (as well as other traits) in older adults. However, the early developmental trajectory of this important social inference has remained largely unexamined. One study reported that children attribute trustworthiness to faces from 3 years of age and these judgments are adult like by 5–6 years of age (Cogsdill, Todorov, Spelke, & Banaji, 2014). Crucially, however, participants in this study were not directly asked to make trustworthiness judgments. Instead, mean/nice evaluations were assumed to approximate trustworthiness judgments. Thus, although the findings from this study provide evidence of basic face-based evaluations of likeability in young children, their relationship to trust judgments is questionable.

The present study is the first to investigate the developmental trajectory of explicit trust judgments of faces. One previous study used a comparable methodology to explore facial trustworthiness attributions in children with autism and noted exciting signs of trust inferences in their typical comparison group, who were aged as young as 6 years (Caulfield, Ewing, Burton, Avard, & Rhodes, 2014). Crucially, this experiment did not compare outcomes between age groups or relate these judgments to those of adults. Nevertheless, based on this result and other evidence that young children form inferences of competence (Antonakis & Dalgas, 2009), propensity towards aggression (Short *et al.*, 2012), and likeability (Cogsdill *et al.*, 2014) from faces, it seems likely that they also form clear impressions of trustworthiness in early childhood. Importantly, however, the maturation of these judgments might be constrained by children's developing sensitivity to facial expressions. Given that emotional expressions strongly influence adults' trustworthiness judgments, developmental improvements in the ability to read expression cues (Durand, Gallay, Seigneuric, Robichon, & Baudouin, 2007; Gao & Maurer, 2009,

2010; Herba & Phillips, 2004) may also contribute to increasingly adult-like impressions of trustworthiness throughout childhood.

The present study aimed to investigate the development of facial trustworthiness judgments. We focused upon the school-age years, when children are increasingly called upon to independently evaluate the trustworthiness of others. Children aged 5, 7, and 10 years, and a group of adults rated the appearance of neutral expression faces previously classified (by adults) as looking trustworthy and untrustworthy. If all three groups of children discriminate between these trustworthy and untrustworthy faces with their ratings like the adults, then this would signal that trust judgments emerge in childhood.

We also investigated whether emotion cues modulate trustworthiness judgments similarly in children and adults. To this end, we had participants make additional trustworthiness ratings of faces explicitly displaying happy and angry expressions, that is, trustworthiness ratings of overt expressions. We used two levels of intensity: 25% and 50% to address the possibility that the less intensely expressed emotions could not be reliably interpreted by the youngest age groups. We looked for evidence of temporal extension, that is, whether overt anger expressions diminished the appearance of trustworthiness in faces and overt happy expressions increased it. We also looked for evidence of emotion overgeneralization from trustworthiness ratings of emotionally neutral (0% happy/angry) faces, that is, whether very subtle angry cues present in emotionally neutral faces diminished the appearance of trustworthiness and very subtle happy cues increased it. Finally, we included an expression recognition measure to determine whether participants showed the expected developmental improvements in expression processing abilities across age groups.

Method

Participants

Participants were 44 adults (17–48 years, $M = 20.8$, $SD = 8.0$; 12 males), 25 children aged 10.1–11.4 years ($M = 10.7$, $SD = 0.4$; 10 males), hereafter referred to as 10-year-old children, 33 children aged 6.1–7.9 years ($M = 7.1$, $SD = 0.6$; 18 males), hereafter referred to as 7-year-old children, and 29 children aged 4.6–5.8 years ($M = 5.4$, $SD = 0.3$; 18 males), hereafter referred to as 5-year-old children. Adult participants were undergraduate psychology students that participated for course credit. Children were recruited from local schools. One 7-year-old child was excluded due to difficulties with concentration.

An additional group of 10 adults (18–33 years, $M = 25.9$, $SD = 5.2$; five males) rated the neutral face stimuli on their subtle resemblance to happy and angry expressions.

Procedure

All parents provided written consent prior to their child's participation in the project, and all children gave verbal and written consent before taking part. The adult participants also provided written consent prior to participation. Participants rated the trustworthiness of neutral faces, then the trustworthiness of expressive faces, and then completed an expression recognition task. This task order was intended to ensure that participants' early judgments were not primed by the emotional cues of the latter tasks.

These tasks were presented to all participants as a game, which took adults approximately 45 min (one session) and children approximately 60 min (two 30-min

sessions) to complete. They were told that an Alien named Zeb needed help to complete a series of missions to learn more about human trustworthiness. On arrival, participants were invited to informally discuss their notions of trustworthiness with the experimenter, before being given a brief description of trustworthiness that focused on three key elements: Honesty, reliability, and emotional trust (Rotenberg, 1994; Rotenberg *et al.*, 2005). They then answered six questions to confirm that they understood this operationalization of trustworthiness, for example, 'Sarah watched her little brother like she promised. Would you trust Sarah?' If participants responded incorrectly to any item, we repeated our description of trust and they were given the six questions once again. No participant required any further explanation to achieve perfect performance.

Trustworthiness from neutral faces

The stimuli for this task were 40 trustworthy and 40 untrustworthy looking Caucasian faces with neutral expressions (20 males in each category) selected from the University of Western Australia Facelab database (Rhodes, Simmons, & Peters, 2005) based on prior adult ratings of trustworthiness (scale ranged from 1 = *not at all trustworthy* to 9 = *extremely trustworthy*; trustworthy set: $M = 6.7$, $SD = 0.5$, untrustworthy set: $M = 3.5$, $SD = 0.4$) (Haynes, 2011). A different group of adult participants rated these stimuli for their resemblance to happy and angry facial expressions. These additional ratings of emotional expressions (happy, angry) were obtained using a 7-point scale (1 = *not at all happy/angry* to 7 = *very happy/angry*). Participants were informed that all faces would be emotionally neutral but nevertheless could show subtle variations in emotional information and that they should still try to use the whole range of the scale, in line with the directions given in another study under similar circumstances (Todorov & Duchaine, 2008).

On each trial, a face was presented on screen for 1,500 ms for participants to rate with the number keys using a 7-point scale consisting of numbered cups (1 = *not very trustworthy* to 7 = *very trustworthy*) (see Cooper, Geldart, Mondloch, & Maurer, 2006). Faces subtended an average visual angle of 8.6° × 6.6° at an approximate viewing distance of 50 cm. Each trial was initiated with a space-bar press and participants made their rating immediately after the face had disappeared from the screen. The 80 faces were presented in randomized order in four blocks of 20 trials. Between each block, participants were given a break in which they were told 'fun facts' about Zeb the alien. Participants began with 16 practice trials: Eight trials using well-known cartoon faces (four trustworthy, four male) and eight trials using real faces (four trustworthy, four male). Here and elsewhere, stimuli used for practice trials were not presented during the main task.

Trustworthiness from expressive faces

This task matched the structure of the above task, except that the to-be-rated stimuli displayed happy and angry, as well as neutral expressions (Figure 1). These stimuli were generated from 20 additional Caucasian faces (10 male), with mid-range trustworthiness ratings ($M = 5.0$, $SD = 0.4$) (also taken from the University of Western Australia Facelab database). Each identity was morphed, using Fantamorph v5.3.1 (http://www.fantamorph.com), with three composite faces displaying happy, angry, and neutral expressions, respectively (each an average of 50 identities, from Skinner & Benton, 2010). For the neutral face condition, we also morphed each original face (neutral expression) image towards a neutral composite. This process ensured that all stimuli presented during

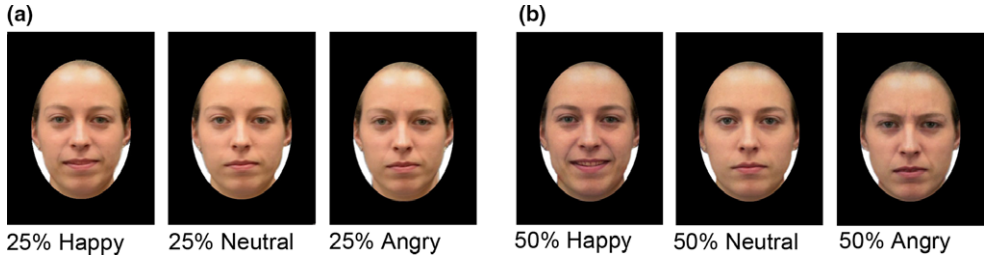

Figure 1. An example of a happy, neutral, and angry expression at 25% intensity (a) and 50% intensity (b). This identity did not appear in the stimulus set.

this task were morphs. Standard morphing procedures were used to create 25% and 50% morphs by blending the original faces with the expression composites in different proportions, for example, 25% angry morph was a 75/25 blend of an original face and the angry composite. Faces subtended an average visual angle of 8.4° × 6.3° at an approximate viewing distance of 50 cm. There were 120 trials in total: 3 expressions (happy, angry, neutral) × 2 intensities (25%, 50%) × 20 identities, which were presented in randomized order in six blocks of 20 trials. Participants began with six practice trials to familiarize them with the stimuli: 3 expressions (happy, angry, neutral) × 2 intensities (25%, 50%).

Expression recognition

The 120 faces from the trustworthiness from expressive faces task were presented again for 1,500 ms each and participants classified each face as happy, angry, or neutral ('when you feel nothing, neither happy or angry') with a labelled key press. The faces were presented in randomized order in four blocks of 30 trials. Participants began with six practice trials: 3 expressions (happy, angry, neutral) × 2 intensities (25%, 50%).

Results

Four extreme scores defined by SPSS[1] (two adult ratings of the expression faces and two adult expression recognition scores) were replaced with the next lowest score. In all but two conditions (adult trustworthiness from neutral faces and adult expression recognition for the 50% angry faces), data were normally distributed. We confirmed our results in the conditions with skewed distributions with nonparametric tests (see below). There were significant violations of the homogeneity of variance assumption for all conditions. Therefore, significant effects were confirmed with Welch's adjusted *F*-ratios and, where appropriate, independent samples *t*-tests are reported with unequal variance assumed (Field, 2009).

Trust ratings of neutral faces

Descriptive statistics for the trust ratings of faces are shown in Table 1. We used the mean difference between ratings of the trustworthy and untrustworthy faces (trustworthy

[1] *An extreme score in SPSS is defined as a value more than three interquartile ranges from the end of the interquartile range box.*

Table 1. Mean (*SD*) for trustworthiness ratings of the trustworthy and untrustworthy faces for each age group

	Trustworthy *M* (*SD*)	Untrustworthy *M* (*SD*)
5-Year-olds	4.4 (0.6)	3.6 (0.9)
7-Year-olds	4.6 (0.7)	3.7 (0.9)
10-Year-olds	4.7 (0.4)	3.0 (0.6)
Adults	4.7 (0.4)	2.9 (0.6)

minus untrustworthy) as our measure of participants' sensitivity to facial trustworthiness (Figure 2). One sample *t*-tests revealed that sensitivity was significantly >0 for all four age groups, all ts > 4.95, ps < .001, rs > .68, which indicated that even the youngest children significantly differentiated between the trustworthy and untrustworthy faces.

A one-way analysis of variance (ANOVA) on trustworthiness sensitivity scores revealed a significant main effect of age group, $F(3, 127) = 16.60$, $p < .001$, *partial* $\eta^2 = .28$, with significantly less differentiation of the trustworthy and untrustworthy faces in the 5- and 7-year-old groups, than in the 10-year-olds and the adults, ts > 3.68, ps < .001, rs > .45 (Figure 2). There was no significant difference between the 5- and 7-year-olds, $t(60) = 0.54$, $p = .59$, $r = .07$, or between the 10-year-olds and adults, $t(67) = 1.08$, $p = .29$, $r = .13$. These results suggest that impressions of trustworthiness may be adult like by 10 years of age.

Inspection of the mean trustworthiness ratings in Table 1 suggests that the age group differences were driven by the 5- and 7-year-olds giving inflated ratings to the untrustworthy faces. This was confirmed by separate one-way ANOVAs for the untrustworthy and trustworthy faces with age group (5-, 7-, 10-year-olds, adults) as a between-participants factor. There was a significant main effect of age group for the untrustworthy faces, $F(3, 127) = 9.59$, $p < .001$, *partial* $\eta^2 = .19$, with the 5- and

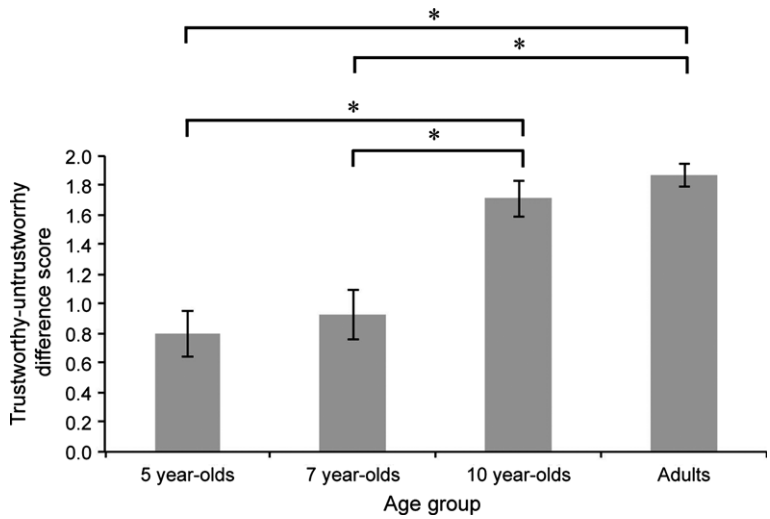

Figure 2. Mean difference (*SEM*) in trustworthiness ratings for the neutral expression trustworthy and untrustworthy faces for each age group. *$p < .05$.

7-year-olds giving significantly more positive ratings to the untrustworthy faces than the 10-year-olds and adults, ts > −2.73, ps < .05, rs > .35, who did not differ from each other, $t(67) = -0.78$, $p = .44$, $r = .09$. For the trustworthy faces, there was no effect of age group, $F(3, 127) = 2.56$, $p = .06$, *partial* $\eta^2 = .06$.

Trustworthiness from expressive faces (temporal extension)

We also investigated how happy and angry expressions modulate trust judgments in children and adults. Given the absence of any specific predictions about the relative magnitude of effects for these two expressions, we did not equate the intensity of our happy and angry stimuli presented during the task and consider them separately in our analysis. In each case, we used a 4 × 2 mixed ANOVA with age group (5-, 7-, 10-year-olds, adults) as a between-participants factor and intensity (25%, 50%) as a repeated-measures factor to examine the modulation of trustworthiness judgments by expression. Descriptive statistics for each age group's trustworthiness ratings of the angry, neutral, and happy expression faces are shown in Table 2.

Angry faces

The dependent variable was the mean difference between trustworthiness ratings of the angry and neutral faces (neutral minus angry), which was significantly >0 for all four age groups for both intensity levels, all ts > 2.32, ps < .05, rs > .40 (Figure 3). This result indicates that anger had a negative effect on perceived facial trustworthiness for all four age groups.

Our ANOVA revealed a significant main effect of age group, $F(3, 127) = 8.32$, $p < .001$, *partial* $\eta^2 = .16$, and intensity, $F(1, 127) = 123.10$, $p < .001$, *partial* $\eta^2 = .49$, with these effects qualified by a significant interaction between age group and intensity, $F(3, 127) = 3.14$, $p = .03$, *partial* $\eta^2 = .07$. Simple tests of main effects showed significant effects of age for the 25% angry expressions, $F(3, 127) = 6.59$, $p < .001$, *partial* $\eta^2 = .14$, and the 50% angry expressions, $F(3, 127) = 7.05$, $p < .001$, *partial* $\eta^2 = .14$. Follow-up independent samples t-tests revealed a similar pattern of age effects for the two intensities (Figure 3). In both intensity conditions, cues to anger influenced trustworthiness ratings significantly more for the adults than for the 5-, 7-, or 10-year-olds, ts > 2.89, ps < .05, rs > .35, who did not differ from each other, ts < 1.75, ps > .09, rs < .24.

Table 2. Mean (*SD*) for trustworthiness ratings of the 25% and 50% angry, neutral, and happy faces for each age group

	25%			50%		
	Angry *M* (*SD*)	Neutral *M* (*SD*)	Happy *M* (*SD*)	Angry *M* (*SD*)	Neutral *M* (*SD*)	Happy *M* (*SD*)
5-Year-olds	3.8 (0.9)	4.1 (0.9)	4.5 (0.9)	3.1 (1.3)	4.2 (0.8)	4.9 (1.0)
7-Year-olds	3.8 (0.8)	4.3 (0.7)	4.6 (0.8)	3.3 (1.3)	4.3 (0.8)	4.8 (1.4)
10-Year-olds	3.2 (0.7)	3.5 (0.7)	4.7 (0.8)	2.4 (0.8)	3.8 (0.6)	5.8 (0.8)
Adults	3.0 (0.6)	3.9 (0.5)	4.6 (0.7)	2.2 (0.5)	4.1 (0.6)	5.5 (0.6)

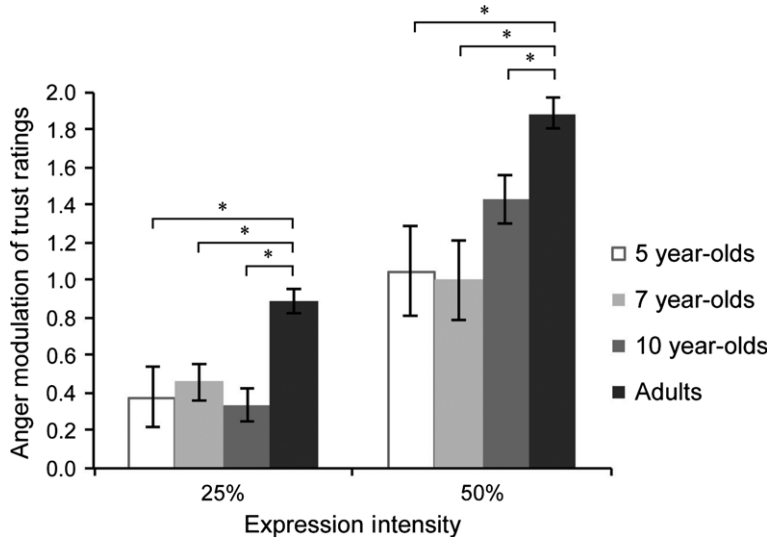

Figure 3. Mean difference (*SEM*) in trustworthiness ratings for angry and neutral expressions at 25% and 50% intensity for each age group. Larger values indicate greater modulation of trust judgments by angry expressions, and zero indicates no modulation. *$p < .05$.

Happy faces

The dependent variable was the mean difference between trustworthiness ratings of the happy and neutral faces (happy minus neutral). One sample *t*-tests indicated that this value was significantly >0 for adults and 10-year-olds at both intensity levels, all *t*s > 7.21, *p*s < .001, *r*s > .79. However, for 7-year-olds, the influence of happy cues was not significant at 25% or 50%, both *t*s < 1.68, *p*s > .10, *r*s < .28. For the 5-year-olds, the influence of happy cues was significant at 50%, $t(28) = 3.26$, $p < .05$, $r = .52$, but not 25%, $t(28) = 1.78$, $p = .09$, $r = .32$.

Our ANOVA revealed a significant main effect of age group, $F(3, 127) = 8.16$, $p < .001$, *partial* $\eta^2 = .16$ (Figure 4). Games–Howell post-hoc tests indicated that happy cues influenced trust judgments significantly more for the 10-year-olds than for the 7- and 5-year-olds, *p*s < .05, and marginally more for the 10-year-olds than the adults, $p = .052$ (see Figure 4). There were no significant differences between the 5-, 7-year-olds, and adults, $p > .05$. There was also a significant main effect of intensity, with the 50% happy cues ($M = 1.1$, $SD = 1.3$) influencing trust judgments more than the 25% happy cues ($M = 0.6$, $SD = 0.9$), $F(1, 127) = 41.84$, $p < .001$, *partial* $\eta^2 = .25$. There was no interaction between intensity and age group, $F(3, 127) = 2.33$, $p = .08$, *partial* $\eta^2 = .05$.

Trustworthiness from neutral faces (emotion overgeneralization)

To assess whether the association between emotional expressions and trust judgments generalized to neutral faces, we looked at the association between perceived trustworthiness and perceived expressions (happy and angry) of the neutral faces.[2] We computed

[2] *Perceived trustworthiness was rated by the 5-, 7-, 10-year-olds, and adults. Perceived expression was rated by the additional group of 10 adults.*

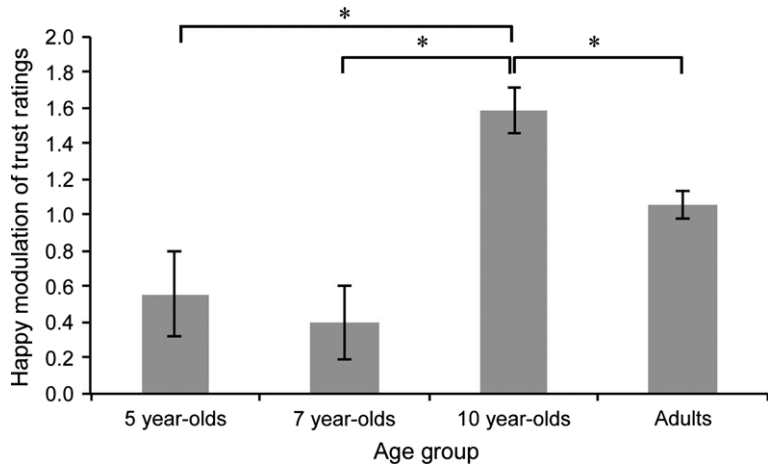

Figure 4. Mean difference (*SEM*) in trustworthiness ratings for the 25% and 50% happy and neutral expressions for each age group. Larger values indicate greater modulation of trust judgments by happy expressions, and zero indicates no modulation. *$p < .05$.

mean ratings (trustworthiness, happiness, and anger) for each face by averaging across the participants' ratings in each age group (see Table 3 for descriptive statistics).

All ratings showed good reliability as measured by Cronbach's alpha (trust ratings in 5-year-olds = .70, 7-year-olds = .78, 10-year-olds = .94, and adults = .98; emotion ratings in additional adults, happiness = .93 and anger = .91).In line with the emotion overgeneralization hypothesis, trustworthiness judgments were significantly positively correlated with perceived happiness and significantly negatively correlated with perceived anger in all four age groups (Table 4). Thus, even with these predominantly neutral expression

Table 3. Descriptive statistics for ratings of trustworthiness, happiness, and anger for the neutral faces

	Trustworthiness					
	5-Year-olds	7-Year-olds	10-Year-olds	Adults	Happiness	Anger
Mean (*SD*)	4.0 (0.7)	4.2 (0.6)	3.9 (1.1)	3.8 (3.8)	3.3 (1.2)	3.3 (1.2)
Range	2.5–4.5	2.9–5.4	1.7–6.5	1.9–5.8	1.3–6.2	1.4–6.0

Table 4. Correlations between trust and emotion ratings of the neutral expression stimuli ($N = 78$) for each age group

	Emotion	
	Happy	Angry
5-Year-olds	.66**	−.67**
7-Year-olds	.73**	−.72**
10-Year-olds	.91**	−.89**
Adults[a]	.80**	−.87**

Note. [a]Spearman's rho corroborates these results; happy: $r = .81$** and angry: $r = -.88$**.
**$p < .001$.

faces, adults and children perceived the stimuli that were subtly happier looking to be more trustworthy and subtly angrier looking to be less trustworthy. Comparison of correlation coefficients between age groups revealed that the 5- and 7-year-olds showed weaker correlations between trust and anger ratings compared to the 10-year-olds and adults, *p*s < .05, who did not differ from each other, $p = .54$. The 5- and 7-year-olds also did not differ from each other, $p = .67$. For the trust and happiness correlations, the 10-year-olds showed stronger effects than the 5-, 7-year-olds, and adults, *p*s < .05, who did not differ from each other, *p*s > .06.

Expression recognition

Hits and false alarm rates were used to calculate signal detection d' scores to measure participants' ability to recognize the 25% and 50% happy and angry expressions.[3] Hit rates of 1 were converted to .975 and false alarm rates of 0 were converted to .0125 following Macmillan and Kaplan's (1985) convention. Once again, we report separate analyses for the angry and happy faces.

Angry faces

Our 4 × 2 mixed ANOVA (as above for the rating tasks) revealed main effects of age group, $F(3, 108) = 28.18$, $p < .001$, *partial* $\eta^2 = .44$, and intensity, $F(1, 108) = 532.37$, $p < .001$, *partial* $\eta^2 = .83$, and a significant interaction between age group and intensity, $F(3, 108) = 3.70$, $p = .01$, *partial* $\eta^2 = .09$. Simple tests of main effects showed significant effects of age for the 25% angry expressions, $F(3, 108) = 30.47$ $p < .001$, *partial* $\eta^2 = .46$, and 50% angry expressions, $F(3, 108) = 18.34$, $p < .001$, *partial* $\eta^2 = .34$. However, again independent samples *t*-tests indicated a similar developmental trend across the two intensities. The 10-year-olds and adults had better expression recognition abilities than the 5- and 7-year-olds for the 25% expressions, all *t*s > 2.19, *p*s < .05, *r*s > .29 (5-year-olds: $M = 0.7$, $SD = 0.\ 6$; 7-year-olds: $M = 0.7$, $SD = 0.5$; 10-year-olds: $M = 1.1$, $SD = 0.6$; adults: $M = 1.9$, $SD = 0.6$), and the 50% expressions, all *t*s > 3.12, *p*s < .05, *r*s > .38 (5-year-olds: $M = 2.2$, $SD = 1.2$; 7-year-olds: $M = 2.4$, $SD = 1.1$; 10-year-olds: $M = 3.2$, $SD = 0.9$; adults: $M = 3.9$, $SD = 0.3$). The adults also had better expression recognition abilities than the 10-year-olds at both intensities, both *t*s > 3.68, *p*s < .05, *r*s > .56. There were no significant differences between the 5- and 7-year-olds for the 25% expressions, $t(60) = -0.18$, $p = .86$, $r = .02$, or 50% expressions, $t(60) = 0.57$, $p = .57$, $r = .07$. Our nonparametric analysis of responses in the 50% angry condition (skewed distribution) corroborated these results, revealing a main effect of age group, $H(3) = 42.44$, $p < .001$ (Kruskal–Wallis test).

Happy faces

The 4 × 2 mixed ANOVA revealed a significant main effect of age group, $F(3, 108) = 15.81$, $p < .001$, *partial* $\eta^2 = .31$. Games–Howell *post-hoc* tests indicated that the adults had better expression recognition abilities than all three groups of children, *p*s < .05 (5-year-olds: $M = 2.1$, $SD = 0.9$; 7-year-olds: $M = 2.3$, $SD = 0.7$; 10-year-olds: $M = 2.8$, $SD = 0.8$; adults: $M = 3.3$, $SD = 0.4$). The 10-year-olds also had better

[3] *Only 25 of the 44 adults were administered the expression recognition task.*

expression recognition abilities than the 5-year-olds, $p = .01$. The 7-year-olds did not significantly differ from the 5- or 10-year-olds, both ps $> .05$. Not surprisingly, there was also a significant main effect of intensity, with better recognition of the 50% expressions ($M = 3.1$, $SD = 1.0$) than the 25% expressions ($M = 2.0$, $SD = 0.9$), $F(1, 108) = 250.67$, $p < .001$, *partial* $\eta^2 = .70$. There was no interaction between intensity and age group, $F(3, 108) = 1.33$, $p = .27$, *partial* $\eta^2 = .04$.

Discussion

This study is the first to demonstrate that children as young as 5 years of age form clear impressions of trustworthiness from faces. We reveal also that these judgments mature throughout childhood, with greater differentiation of trustworthy and untrustworthy faces by 10-year-old children and adults, than by 5- and 7-year-old children. Moreover, we found that 5- and 7-year-old children tend to overestimate the trustworthiness of untrustworthy faces, relative to 10-year-old children and adults. These results suggest that the ability to evaluate the trustworthiness of faces emerges in childhood, but may not be adult like until 10 years of age.

Our results indicate that trust judgments in children, like those in adults, are powerfully influenced by emotional expression cues. Overtly angry expressions diminished the appearance of trustworthiness in faces significantly for 5-, 7-, 10-year-olds, and adults. In addition, overtly happy expressions increased the appearance of trustworthiness, although this effect did not consistently emerge until later in childhood, that is, 10 years of age. This modulation of trust judgments by expression supports the notion that trait attributions in children, like adults, partially reflect temporal extension of typically transient facial cues of emotion to the more enduring characteristic of trustworthiness.

We also observed emotion overgeneralization in children as young as 5 years of age. All four age groups judged those neutral faces displaying very subtle anger cues to appear relatively untrustworthy and those faces displaying very subtle happy cues as relatively trustworthy. Here, stronger effects in the 10-year-olds and adults, relative to the 5- and 7-year-olds, might reflect age-related differences in the reliability of each group's trustworthiness judgments. Regardless, this systematic association between the trustworthiness and perceived emotions of neutral expression faces paralleled our findings with the overt happy and angry facial expressions. These findings indicate that children's trustworthiness judgments, like adults, are not only modulated by obvious expressions of emotion (temporal extension), but also those more subtle emotion cues present in neutral faces (emotion overgeneralization).

Our results suggest that the modulatory power of emotion cues may increase across development. For example, overt angry expressions influenced trust judgments more strongly in the adults than in the children. Similarly, overt happy expressions had a stronger effect for 10-year-olds than 5- and 7-year-olds. Importantly, these increases in the influence of facial expressions were accompanied by age-related increases in participants' sensitivity to the two expressions. Thus, it appears that developmental differences in the influence of happy and angry facial expressions may be, in part, related to developmental differences in our participants' ability to read these cues.

Critically, differences in sensitivity to expressions cannot explain all of the current results. Overt happy cues did not influence trust judgments in 7-year-olds, even though these children showed excellent recognition performance for this expression at both 25%

and 50% intensity. Similarly, the 25% happy expressions did not significantly influence 5-year-olds' trust judgments, despite sound recognition of this expression in this group. One alternative explanation is that weaker effects in these younger age groups may reflect greater variability in responding rather than specific immaturity of trust perception. We carefully designed developmentally appropriate tasks for the current study and worked to ensure that children were motivated to attend closely to stimuli. Nevertheless, it remains possible that age-related differences in general cognitive factors, such as attention, may have affected the 5- and 7-year-olds' performance (Crookes & McKone, 2009; Gilchrist & McKone, 2003; Mondloch, Maurer, & Ahola, 2006). Still, given that much subtler cues to happiness in neutral faces did modulate trustworthiness judgments for the 5- and 7-year-old children, this explanation seems unlikely.

Another possibility is that there may be differences in the comparative influence of overt happy and angry expressions on trust judgments. Our happy and angry stimuli were not equated for perceived emotional intensity, so we could not compare the relative modulatory power of happiness and anger cues. This design choice reflected our focus on whether these expressions modulated trustworthiness judgments in children like adults. However, given the adaptive significance of being able to quickly identify an untrustworthy (and potentially dangerous) conspecific, overt angry expressions might actually modulate trustworthiness judgments earlier and/or more powerfully than overt happy expressions. Future studies might consider equating the perceptual intensity of these expressions to facilitate such comparisons.

There were also other indications that developmental changes in trust judgments might reflect more than just age-related changes in expression sensitivity. For example, although sensitivity to the happy expressions increased with age, these cues had their strongest effects on the trust judgments of 10-year-olds, rather than the adults. We speculate that although the happy expressions increased the perceived trustworthiness of faces for both 10-year-olds and adults, the adults may have had more of an understanding that these expressions were posed, and thus a potentially misleading signal of personality characteristics, than the 10-year-old children. Although school-aged children have been shown to be able to detect differences between genuine and posed smiles, sensitivity is not high (Gosselin, Perron, & Maassarani, 2010). This ability improves across development, particularly between late childhood and adulthood (Gosselin, Perron, Legault, & Campanella, 2002). Therefore, the adult participants in the current study were likely more sensitive to facial markers indicating the genuineness of a smile and may have 'toned down' their initial positive impressions of the faces accordingly.

Our results provide a useful starting point for further investigations of the development of facial trustworthiness judgments. One important question will be how early these judgments emerge. Here, we provide evidence that children as young as 5 years of age form impressions of trustworthiness from faces. However, the socialization of trust begins in the first few years of life, developing with accumulated experience and interpersonal interactions throughout early childhood (Krebs & Hesteren, 1994; Rotenberg, 1995). Thus, it seems possible that appearance-based trust judgments may emerge even earlier in development than reported here.

The current findings revealed some age-related differences in perceptions of trustworthiness from adult faces. It will be interesting to see how these results might generalize to children's perceptions of trustworthiness in own-age peers. There is some evidence that children recognize own-age faces more accurately than other-age faces (Anastasi & Rhodes, 2005). Thus, we might predict more mature-looking perceptions of

trustworthiness and sensitivity to facial expressions when children make trust judgments about the faces of their peers.

It will also be interesting to further examine trust inferences in children with autism spectrum disorder (ASD), who demonstrate social cognitive difficulties (Baron-Cohen *et al.*, 2000; Webb, Faja, & Dawson, 2011). Individuals with ASD often have difficulties extracting emotional information from faces (for reviews see Gaigg, 2012; Harms, Martin, & Wallace, 2010; Nuske, Vivanti, & Dissanayake, 2013; Uljarevic & Hamilton, 2013). Therefore, it seems plausible that expression processing impairments may contribute to atypicalities in the perception of trust. Interestingly, a recent study found no differences in trustworthiness judgments between typically developing children and children with ASD (Caulfield *et al.*, 2014). As with the current study, trustworthiness judgments for both groups of children were influenced by overt happy and angry expressions as well as subtle emotional cues in neutral faces. However, a wide range of ages was collapsed (ages ranged from 6 to 11 years for the typically developing children and 6–12 years for the children with ASD), which could have masked developmental differences between these two groups of children. Given that the current study has demonstrated significant age-related changes in trustworthiness judgments over this age range, the developmental trajectory of trustworthiness judgments in children with ASD may yet be shown to differ from that of typically developing children.

In conclusion, we have demonstrated that children as young as 5 years of age draw inferences about trustworthiness from facial appearances, which become increasingly adult like with age. Furthermore, similar expression cues influence these judgments in children and adults. Overt expressions (temporal extension), as well as subtle expression cues perceptible in neutral faces (emotion overgeneralization), modulated trustworthiness judgments in children and adults. Together with recent evidence of reliable competence inferences from faces by children (Antonakis & Dalgas, 2009), these findings highlight an early-emerging capacity for sophisticated social cognition in young children.

Acknowledgements

This research was supported by the Australian Research Council Centre of Excellence in Cognition and its Disorders (CE110001021), an ARC Professorial Fellowship to Rhodes (DP0877379), and an ARC Discovery Outstanding Researcher Award to Rhodes (DP130102300). We thank Sarah Haynes for supplying the pre-rated trustworthy and untrustworthy stimuli and Andy Skinner and Christopher Benton for providing us with their composite expression stimuli. We also thank Ainsley Read for assistance with recruitment and testing and Eleni Avard for assistance with testing. Ethical approval was granted by the Human Research Ethics Committee of the University of Western Australia.

References

Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review*, *12*, 1043–1047. doi:10.3758/BF03206441

Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play!. *Science*, *323*, 1183–1184. doi:10.1126/science.1167748

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. R. (2000). The amygdala theory of autism. *Neuroscience & Biobehavioral Reviews*, *24*, 355–364. doi:10.1016/S0149-7634(00)00011-7

Caulfield, F., Ewing, L., Burton, N., Avard, E., & Rhodes, G. (2014). Facial trustworthiness judgments in children with ASD are modulated by happy and angry emotional cues. *PLoS ONE*, *9*, e97644. doi:10.1371/journal.pone.0097644

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*, 87–105. doi:10.1016/j.cogpsych.2010.03.001

Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from faces. A developmental study. *Psychological Science*, *25*, 1132–1139. doi:10.1177/0956797614523297

Cooper, P. A., Geldart, S., Mondloch, C. J., & Maurer, D. (2006). Developmental changes in perceptions of attractiveness: A role of experience? *Developmental Science*, *9*, 530–543. doi:10.1111/j.1467-7687.2006.00520.x

Crookes, K., & McKone, E. (2009). Early maturity of face recognition: No childhood development of holistic processing, novel face encoding, or face-space. *Cognition*, *111*, 219–247. doi:10.1016/j.cognition.2009.02.004

Durand, K., Gallay, M., Seigneuric, A., Robichon, F., & Baudouin, J. Y. (2007). The development of facial emotion recognition: The role of configural information. *Journal of Experimental Child Psychology*, *97*(1), 14–27. doi:10.1016/j.jecp.2006.12.001

Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2014). Perceived trustworthiness of faces drives trust behaviour in children. *Developmental Science*, *18*, 327–334. doi:10.1111/desc.12218

Field, A. (2009). *Discovering statistics using SPSS*. London, UK: Sage.

Franklin, R., & Zebrowitz, L. A. (2013). Older adults' trait impressions of faces are sensitive to subtle resemblance to emotions. *Journal of Nonverbal Behavior*, *37*, 139–151. doi:10.1007/s10919-013-0150-4

Gaigg, S. B. (2012). The interplay between emotion and cognition in Autism Spectrum Disorder: Implications for developmental theory. *Frontiers in Integrative Neuroscience*, *6*, 1–35. doi:10.3389/fnint.2012.00113

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*, 503–521. doi:10.1016/j.jecp.2008.11.002

Gao, X., & Maurer, D. (2010). A happy story: Developmental changes in children's sensitivity to facial expressions of varying intensities. *Journal of Experimental Child Psychology*, *107*, 67–86. doi:10.1016/j.jecp.2010.05.003

Gilchrist, A., & McKone, E. (2003). Early maturity of face processing in children: Local and relational distinctiveness effects in 7-year-olds. *Visual Cognition*, *10*, 769–793. doi:10.1080/13506280344000022

Gosselin, P., Perron, M., Legault, M., & Campanella, P. (2002). Children's and adults' knowledge of the distinction between enjoyment and nonenjoyment smiles. *Journal of Nonverbal Behavior*, *26*, 83–108.

Gosselin, P., Perron, M., & Maassarani, R. (2010). Children's ability to distinguish between enjoyment and non-enjoyment smiles. *Infant and Child Development*, *19*, 297–312. doi:10.1002/icd.648

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, *20*, 290–322. doi:10.1007/s11065-010-9138-6

Haynes, S. (2011). *Judgments of trustworthiness from faces: Do children and adults judge alike?* (Psychology Honours Dissertation). The University of Western Australia.

Herba, C., & Phillips, M. (2004). Annotation: Development of facial expression recognition from childhood to adolescence: Behavioural and neurological perspectives. *Journal of Child Psychology and Psychiatry*, *45*, 1185–1198. doi:10.1111/j.1469-7610.2004.00316.x

Krebs, D. L., & Hesteren, F. V. (1994). The development of altruism: Toward an integrative model. *Developmental Review*, *14*, 103–158. doi:10.1006/drev.1994.1006

Lavater, J. C. (1789). *Essays on physiognomy*. London, UK: William Tegg and Co.

Macmillan, N., & Kaplan, H. (1985). Detection theory analysis of group data: Estimating sensitivity from the average hit and false-alarm rates. *Psychological Bulletin*, *98*, 185–199. doi:10.1037/0033-2909.98.1.185

McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, *90*, 215–238. doi:10.1037/0033-295X.90.3.215

Mondloch, C. J., Maurer, D., & Ahola, S. (2006). Becoming a face expert. *Psychological Science*, *17*, 930–934. doi:10.1111/j.1467-9280.2006.01806.x

Nuske, H. J., Vivanti, G., & Dissanayake, C. (2013). Are emotion impairments unique to, universal, or specific in autism spectrum disorder? A comprehensive review. *Cognition & Emotion*, *27*, 1042–1061. doi:10.1080/02699931.2012.762900

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

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, *9*(1), 128–133. doi:10.1037/a0014520

Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable facial configurations affect strategic choices in trust games with or without information about past behavior. *PLoS ONE*, *7*, e34293. doi:10.1371/journal.pone.0034293

Rhodes, G., Simmons, L. W., & Peters, M. (2005). Attractiveness and sexual behavior: Does attractiveness enhance mating success? *Evolution and Human Behavior*, *26*, 186–201. doi:10.1016/j.evolhumbehav.2004.08.014

Rotenberg, K. J. (1994). Loneliness and interpersonal trust. *Journal of Social and Clinical Psychology*, *13*, 152–173. doi:10.1521/jscp.1994.13.2.152

Rotenberg, K. J. (1995). The socialisation of trust: Parents' and children's interpersonal trust. *International Journal of Behavioral Development*, *18*, 713–726. doi:10.1177/016502549501800408

Rotenberg, K. J., Fox, C., Green, S., Ruderman, L., Slater, K., Stevens, K., & Carlo, G. (2005). Construction and validation of a children's interpersonal trust belief scale. *The British Journal of Developmental Psychology*, *23*, 271–293. doi:10.1348/026151005X26192

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

Secord, P. (1958). Facial features and inference processes in interpersonal perception. In R. Tagiuri & P. Luigi (Eds.), *Person perception and interpersonal behavior* (pp. 300–315). Standford, CA: Standford University Press.

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

Skinner, A., & Benton, C. (2010). Anti-expression after effects reveal prototype-referenced coding of facial expressions. *Psychological Science*, *21*, 1248–1253. doi:10.1177/0956797610380702

Sutherland, C., Oldmeadow, J., Santos, I., Towler, J., Burt, D., & Young, A. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*, 105–118. doi:10.1016/j.cognition.2012.12.001

Todorov, A., & Duchaine, B. (2008). Reading trustworthiness in faces without recognizing faces. *Cognitive Neuropsychology*, *25*, 395–410. doi:10.1080/02643290802044996

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

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders*, *43*(7), 1–10. doi:10.1007/s10803-012-1695-5

Webb, S. J., Faja, S., & Dawson, G. (2011). Face processing in autism. In A. J. Calder, G. Rhodes, M. H. Johnson & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 839–856). Oxford, UK: Oxford University Press.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100 ms exposure to a face. *Psychological Science*, *17*, 592–598. doi:10.1111/j.1467-9280.2006.01750.x

Zebrowitz, L. A. (1996). Physical appearance as a basis for stereotyping. In M. H. N. McRae & C. Stangor (Eds.), *Foundation of stereotypes and stereotyping* (pp. 79–120). New York, NY: Guildford Press.

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

Zebrowitz, L. A., & Collins, M. A. (1997). Accurate social perception at zero acquaintance: The affordances of a Gibsonian approach. *Personality and Social Psychology Review*, *1*, 204–223. doi:10.1207/s15327957pspr0103_2

Zebrowitz, L. A., Kikuchi, M., & Fellous, J. (2010). Facial resemblance to emotions: Group differences, impression effects, and race stereotypes. *Journal of Personality and Social Psychology*, *98*, 175–189. doi:10.1207/s15327957pspr0103_2

Received 17 February 2015; revised version received 14 September 2015