
Article:
Strachan, James, Kirkham, Alexander James, Manssuer, Luis et al. (2017) Incidental learning of trust from eye-gaze: Effects of race and facial trustworthiness. VISUAL COGNITION. pp. 802-814. ISSN: 1350-6285

https://doi.org/10.1080/13506285.2017.1338321

Incidental learning of trust from eye-gaze:

Effects of race and facial trustworthiness

James W. A. Strachan[1,3], Alexander J. Kirkham[1], Luis R. Manssuer[2], Harriet Over[1], and Steven P. Tipper[1]

1. University of York, Heslington
2. University of Sheffield
3. Central European University, Budapest

Corresponding author:

James W.A. Strachan,

Department of Cognitive Science,

Central European University,

Oktober 6 utca 7,

Budapest 1051,

Hungary.

Contact email: james.wa.strachan@gmail.com

Data files are available for download from the following DOI:

10.17605/OSF.IO/P8Q7Z

Abstract

Humans rapidly make inferences about individuals' trustworthiness on the basis of their facial features and perceived group membership. We examine whether incidental learning about trust from shifts in gaze direction is influenced by these facial features. To do so, we examined two types of face category: the race of the face and the initial trustworthiness of the face based on physical appearance. We find that cueing of attention by eye-gaze is unaffected by race or initial levels of trust, whereas incidental learning of trust from gaze behaviour is selectively influenced. That is, learning of trust is reduced for other race faces, as predicted by reduced abilities to identify members of other races (Experiment 1). In contrast, converging findings from an independently gathered set of data showed that the initial trustworthiness of faces did not influence learning of trust (Experiment 2). These results show that learning about the behaviour of other-race faces is poorer than for own-race faces, but that this cannot be explained by differences in the perceived trustworthiness of different groups.

Keywords: gaze cueing; trustworthiness; own-race bias; incidental social learning; face perception

Incidental learning of trust from eye-gaze:

Effects of race and facial trustworthiness

With just a brief glance at a person, the basic physiognomy of the face provides sufficient information for its rapid classification. In a very short space of time (between 100-200ms), people make inferences about whether a face belongs to their own or another social group and whether a person can or cannot be trusted (e.g., Caldara, Rossion, Bovet, & Hauert, 2004; Willis & Todorov, 2006). However, basing our behavioural decisions on such physiognomy can be harmful, leading to rigid responses that reinforce discrimination. In addition, it can render us incapable of responding to the dynamics of behaviour and adjusting our behaviour towards others in light of incoming information. The central focus of this paper is to investigate the relationship between face physiognomy that enables a rapid classification of a face in terms of race or levels of trust, and how the subtle dynamic changing behaviours that signal deception can be learned.

Previous work has indeed shown that changeable aspects of faces, such as view direction or emotion, can alter judgments of trustworthiness. For example, faces expressing a smile are trusted more than faces expressing anger (e.g., Caulfield, Ewing, Burton, Avard, & Rhodes, 2014; Sutherland, Young, & Rhodes, 2016). Importantly, within such studies these properties of a face are present, and therefore attended to whilst trustworthiness judgments are being made. However, there are subtle dynamic facial behaviours that could be learned, even when they are not explicitly attended to. These behaviours might affect later trustworthiness judgments even though the face does not possess these properties at a later judgment time. To investigate this issue, we explore the incidental learning of patterns of gaze shifts made by a face that is being ignored.

When we see the eyes of a face re-fixate to another point in space our own attention follows their gaze. This results in enhanced processing for objects in that area of space, and poorer or disrupted processing for objects elsewhere (Friesen & Kingstone, 1998; for review, see Frischen, Bayliss, & Tipper, 2007). Previous research has shown that this gaze cueing effect is powerful and difficult to inhibit, that people show cueing effects even when they know that gaze direction is uninformative or that the likelihood of invalid cues is high (Driver et al., 1999), and that they are sensitive to misleading gaze cues even when they know the true future location of the target (Galfano et al., 2012).

Given that these cues are such a salient and powerful form of nonverbal communication, they can be used to direct an interaction partner's attention around the environment, either helpfully (directing them to appetitive or aversive stimuli that merit attention) or deceitfully (misdirecting them away from such stimuli). Furthermore, observers are able to detect consistent gaze patterns, and use this information to guide subsequent decisions: faces that provide consistently valid (helpful) cues are both rated and treated as more trustworthy than those that provide consistently invalid (deceitful) cues (Bayliss & Tipper, 2006; Manssuer, Pawling, Hayes, & Tipper, 2016; Manssuer, Roberts, & Tipper, 2015; Rogers et al., 2014; Strachan, Kirkham, Manssuer, & Tipper, 2016; Strachan & Tipper, 2017).

This paradigm, which measures incidental learning of trust in that participants are instructed to ignore the faces during gaze cueing, has been used multiple times to explore different facets of this trust learning process. The effect has been shown to be durable (Strachan & Tipper, 2017), related to involuntary facial muscular and electro-cortical activity associated with emotional responses

(Manssuer et al., 2016), and sensitive to the emotion of the cueing face (Bayliss, Griffiths, & Tipper, 2009; Strachan et al., 2016). However, one outstanding question is how a fixed property of the cueing face that conveys higher order social information, such as its race, may affect (or not affect) this learning.

Racial group membership

Humans are social creatures, and we rely on social groups in order to survive. These social groups can range from a smaller personal scale (e.g. a close circle of friends) to a much larger societal scale (our national identity, gender, race, etc.; Lickel, Hamilton, & Sherman, 2001). People prefer individuals who are members of their in-group over their out-group, even when this group distinction is a new category that has been learned in the laboratory (Allen & Wilder, 1975; Tajfel, Billig, Bundy, & Flament, 1971).

One of the most studied and historically important social group categories is that of race. Aside from showing preferential biases towards own-race over other-race faces (Dasgupta, Mcghee, Greenwald, & Banaji, 2000), people are also better at recognising the posed emotions of own-race than other-race faces (Elfenbein & Ambady, 2002a, 2002b), remember own-race faces better than other-race faces (Meissner & Brigham, 2001), and are more sensitive to gaze cues provided by own- than other-race faces (Dalmaso, Galfano, & Castelli, 2015; Pavan, Dalmaso, Galfano, & Castelli, 2011). There is also evidence that these own-race biases are linked to decisions about trustworthiness in both explicit ratings and economic games (Stanley, Sokol-Hessner, Banaji, & Phelps, 2011).

In Experiment 1 we explore whether trust learning is the same for individuals who differ in terms of racial group membership. To this end, we test British Caucasian participants using images of White (own race) and East Asian

(other race) faces in the cueing paradigm and ask participants to rate them in terms of trust at the beginning and end of the experiment.

We predict that learning about other-race faces from subtle and incidental gaze cues will be impaired. The trust learning task requires that gaze patterns be associated with a specific face identity. That is, when a particular face produces consistent gaze behaviour, for example always looking away from targets, this behaviour has to be associated with the cognitive representation of that individual. However, if as previous research suggests (Dalmaso et al, 2015), people can be less susceptible to gaze cues provided by other-race faces, and less efficient at remembering other-race faces (Meissner & Brigham, 2001; Sessa & Dalmaso, 2016), it stands to reason that learning for these faces would also be impaired. Therefore we expect that incidental learning of trust will be weaker in other as compared to own race faces.

Experiment 1

This experiment explored whether the original trust learning effect (Bayliss & Tipper, 2006) emerges similarly for faces that belong to real world social in- and out-groups by using White and East Asian faces with Caucasian British participants.

Methods

Participants.

In Experiment 1, there were 30 participants in total (all Caucasian British, 26 female, M_{age} 22.23). Sample size was decided on the basis of previous studies that investigate incidental social learning from gaze cues, which use between 20 and 30 participants (Manssuer et al., 2016, Experiment 2; Rogers et al., 2014; Strachan et al., 2016; Strachan & Tipper, 2017). Given that our manipulation

involved a more complicated design than previous experiments (with race as an additional independent variable) we opted for the larger end of this range with 30 participants. All participants provided written consent and the study was given ethical approval by the Ethics Committee of the University of York Psychology Department.

Stimuli.

Target stimuli for the object categorisation task were kitchen and garage object images used in Bayliss and Tipper (2006). There were 13 unique objects in each category (kitchen/garage) and these appeared in both left and right orientations. All stimuli were coloured in blue. In total there were 52 individual images used in the experiment.

Face stimuli throughout the experiment were taken from the MR2 Face Database (Strohminger et al., 2015). This database comes with a set of ratings for each face on a range of attributes, including trustworthiness on a scale of 1 (untrustworthy) to 7 (trustworthy); these are publicly available on the Open Science Framework (OSF; https://osf.io/uwk4v/). The 8 (4 female, 4 male) East Asian faces and 8 (4 female, 4 male) White faces were selected on the basis of these ratings to be similar in terms of apparent trustworthiness (East Asian faces: $M = 4.13$, *s.d.* = 0.14; White faces: $M = 4.12$, *s.d.* = 0.21). These images were then edited in Adobe Photoshop CS6 to remove the grey background and edit the direction of eye gaze to create three versions of each face: straight, left, and right gaze. These stimuli were used in the gaze-cueing procedure, while faces with unedited eyes were used in the trustworthiness ratings and one-back procedure.

For each participant, the cueing behaviour of faces was set such that each face would provide a valid or invalid cue 100% of the time, and face validity was

manipulated orthogonally to race, such that there were four conditions of faces: East Asian Valid, East Asian Invalid, White Valid, and White Invalid, with four identities in each condition (see Figure 1a). The validity of faces was counterbalanced across participants.

[Figure 1 approximately here]

The study was run on an Intel Core i5 PC with a 21.5" monitor. The experiment was presented using E-Prime 2.0 software with a white background throughout and the resolution set to 1024x768 pixels. Participants were sat approximately 60cm from the display, and during trustworthiness ratings the face stimuli had a visual angle of 19.29˚ horizontally and 20.97˚ vertically, while during gaze-cueing the face stimuli had a visual angle of 13.36˚ horizontally and 14.93˚ vertically.

Design and procedure.

Participants initially completed a one-back recognition task at the beginning of the experiment with all the faces used later as stimuli. This was done because Strachan and Tipper (2016) have shown that greater familiarity with faces can improve trust learning. In that experiment, participants were asked to match faces across images that varied in expression and viewpoint, but the MR2 database does not provide any such variation. Therefore in this experiment we included a one-back recognition task, where faces were presented in sequence and participants had to respond with the SPACE bar if they saw the same face repeated twice in a row. This encourages participants to encode details of the faces and store them in working memory, at least until the next face is shown, and with repeated exposures this should allow participants to become familiar with the face identities.

Participants were told that they would be asked to perform an object categorisation task on images of objects that appeared on the left or right side of the screen, and to respond with whether these were garage or kitchen objects. They were also told that the central face images were irrelevant and to be ignored. Before the experiment, participants were allowed to study printed versions of the kitchen/garage items, in order to familiarise themselves. This was done firstly to ensure that participants knew what each object was, and secondly to ensure that early responses from the first trial block were not confounded by uncertainty as to the object categories of the targets.

Each trial began with a 600ms fixation cross in the centre of the screen, which was then replaced by a face showing a direct gaze for 1,500ms. The face then shifted gaze either to the left or the right for 500ms before the target stimulus appeared on either the same (valid) or opposite (invalid) side of the gaze direction. The target stimulus remained either until the participant's response was logged or until 2,500ms had passed, following which participants received feedback from an error tone that would sound if an incorrect response were logged. The face then shifted back to direct gaze for another 1,000ms. A blank screen followed for 500ms before the next trial began. The trial structure is shown in Figure 1b.

The object categorisation responses were the H key and the space bar of a keyboard, chosen because the H key appears directly above the space bar on QWERTY keyboards and this direction was orthogonal to the possible location of the target. Participants were instructed to respond with their index finger on the H key and thumb on the space bar. For half the participants, H represented kitchen objects, while for the other half it represented garage objects.

In total, there were five blocks of 32 trials each, and each face appeared twice in each block, once gazing left and once right (10 times in total across the experiment; five left, five right, but always either valid or invalid depending on the identity). The order of faces was randomised, as were the order of target objects, the side that the target appeared, and the order of valid and invalid trials.

At the beginning and the end of the experiment, participants were shown all 16 faces (un-manipulated original images) in a random order and asked to rate how trustworthy they found them. A calibration screen would appear with the question, "How TRUSTWORTHY do you think this person is?" with the word 'START' written beneath. Participants had to click the word 'START' with the mouse to progress the trial, after which the face would appear for 1,000ms. Following this, the face would disappear and a screen with an uninterrupted rating scale appeared. Participants were instructed to click along the scale with the computer mouse at the point that corresponded to how trustworthy they thought the person was. The scale recorded responses between -100 and +100, calculated by the distance from the centre of the line of the participants' mouse click – responses to the left of the centre of the line were coded as negative, while those to the right were coded as positive (these were indicated on the screen with a – and + sign at either end of the scale). ▪

Data analysis

Before the data were analysed, participants' responses were filtered to remove all error trials (where participants reported the incorrect answer) and reaction time (RT) outliers – RTs below 250ms (too short to process the stimuli; 0.10% of trials) and above 2,500ms (indicating that participants had not given a

response in the allotted time; 0.44% of trials). The number of remaining trials was then compared with the original number of trials to check that all participants retained at least 70% of their total trials and had not scored below 70% total correct on any one condition. No participants were removed on this basis.

As well as RT filters, we also examined participants' pre-ratings. Participants' ratings to in-group and out-group faces were averaged and examined to ensure that the average for neither group exceeded 70 on the 100-point scale in either direction. This was done because an average to one group that exceeded 70 suggested that participants gave ratings to multiple faces that used the far ends of the scale before any trustworthiness induction was performed, resulting in a floor or ceiling effect where any effect of our manipulation would be masked. This is to say, for example, if one participant rated other-race faces as appearing extremely untrustworthy (using the far left of the trustworthiness scale) they would be excluded as any trust learning would be subject to a floor effect. No participant was removed on this basis in Experiment 1.

Gaze cueing was examined in terms of reaction times (RTs) and accuracy rates separately using 2x2 repeated measures ANOVAs, with race (own/other race) and validity (valid/invalid) as factors. Incidental trust learning was tested with a 2x2x2 repeated measures ANOVA with time (before/after the experiment), validity and race as repeated measures factors and trustworthiness rating as dependent variable. All analysis was run using the *ez* package in the statistical software *R*.

Results

Gaze cueing

[Figure 2 approximately here]

The RT results of Experiment 1 are shown in Figure 2 (top row). A 2x2 repeated measures ANOVA of race and validity on mean RTs found a main effect of validity ($F(1,29) = 17.84$, $p < .001$, $\eta^2_P = 0.38$) where responses were faster on validly-cued trials, but none of race ($F(1,29) = 1.93$, $p = 0.175$, $\eta^2_P = 0.06$) and no interaction of the two ($F(1,29) = 0.25$, $p = 0.622$, $\eta^2_P = 0.01$).

The accuracy results (calculated as percent correct) of Experiment 1 are shown in Table 1. A 2x2 ANOVA of race and validity found no effect of validity ($F(1,29) = 0.19$, $p = 0.664$, $\eta^2_P = 0.01$), or of race ($F(1,29) = 0.01$, $p = 0.906$, $\eta^2_P = 0.00$), and no interaction ($F(1,29) = 1.02$, $p = 0.320$, $\eta^2_p = 0.03$).

[Table 1 approximately here]

Trustworthiness Ratings

Trust Learning. The results of trustworthiness ratings at the beginning and end of Experiment 1 are shown in Figure 3 (top row). A 2x2x2 ANOVA looking at time, validity and race found a main effect of time ($F(1,29) = 12.29$, $p = 0.001$, $\eta^2_p = 0.30$). There was also a main effect of validity ($F(1,29) = 5.18$, $p = 0.030$, $\eta^2_P = 0.15$), as invalid faces were rated as less trustworthy than valid, but none of race ($F(1,29) = 0.64$, $p = 0.430$, $\eta^2_P = 0.02$), as own-race faces were not rated as more trustworthy over the course of the whole experiment than other-race faces. A significant interaction of time and validity was found ($F(1,29) = 9.07$, $p = 0.005$, $\eta^2_P = 0.24$), indicating that there was significant learning of trust over time as a function of gaze cueing behaviour. There was also a significant interaction of time and race ($F(1,29) = 4.84$, $p = 0.036$, $\eta^2_P = 0.14$), which appears

to be driven by the fact that a slight own-race bias in pre-ratings was less evident in post-experiment ratings, and a non-significant interaction of validity and race ($F(1,29) = 3.41$, $p = 0.075$, $\eta^2_P = 0.11$). Importantly, there was a three-way interaction of time, validity, and race, indicating that trust learning over the experiment was affected by race ($F(1,29) = 4.45$, $p = 0.044$, $\eta^2_P = 0.13$).

[Figure 3 approximately here]

We broke this down into separate 2-way ANOVAs that looked at the effects of validity and race at the beginning and the end of the experiment, separately. At the beginning of the experiment, there was no main effect of validity ($F(1,29) = 0.68$, $p = 0.415$, $\eta^2_P = 0.02$), as participants had not been exposed to faces' valid or invalid behaviours at this point. Participants did rate own-race faces on average as more trustworthy ($M = 9.89$, *s.d.* = 30.53) than other-race faces ($M = 4.73$, *s.d.* = 25.55) but this was not significant ($F(1,29) = 3.41$, $p = 0.075$, $\eta^2_P = 0.11$) . There was no interaction of validity and race ($F(1,29) = 0.02$, $p = 0.884$, $\eta^2_P = 0.00$).

On the other hand, at the end of the experiment there was a main effect of validity ($F(1,29) = 7.84$, $p = 0.009$, $\eta^2_P = 0.21$), and although the main effect of race was not significant ($F(1,29) = 0.04$, $p = 0.842$, $\eta^2_P = 0.00$) there was a significant interaction between the two ($F(1,29) = 8.71$, $p = 0.006$, $\eta^2_P = 0.23$),[1] confirming greater trust learning (valid-invalid) for in-group faces than out-group. At the end of the experiment there were significant differences between valid ($M = 12.81$, *s.d.* = 19.19) and invalid faces for own-race faces ($M = -9.56$, *s.d.* = 28.15; $t(29) = 3.20$, 95%CI [8.08, 36.66], $p = .003$) but trust ratings were not significantly different for other-race faces (valid: $M = 7.57$, *s.d.* = 20.75; invalid: *M*

= -2.96, *s.d.* = 20.75; $t(29) = 1.98$, 95%CI [-0.33, 21.38], $p = .057$; Bonferroni-corrected α: .025).

There are two key findings from Experiment 1. First, gaze cueing where participants follow the gaze direction of another person is unaffected by whether the viewed face is a racial in-group or out-group member. This is surprising given previous research that shows that race can affect susceptibility to gaze cues (Chen et al., 2017; Chen & Zhao, 2015; Dalmaso et al., 2015; Pavan et al., 2011). However, some of these findings suggest that this effect of race is mediated by a sense of inter-group threat – that is, when participants feel that out-group faces appear threatening (due to their out-group status), people are less susceptible to gaze cues. Other studies that have found effects of race often use Black faces (rather than East Asian faces), which for White participants often carry a threatening connotation. In contrast, stereotype content for East Asian identities tends to be more nuanced (Lin, Kwan, Cheung, & Fiske, 2005), and may be less likely to be spontaneously perceived as threatening. Indeed, evidence from EEG suggests that race may affect face processing more when faces show direct gaze, rather than averted (Sessa & Dalmaso, 2016), which suggests that some additional context is required for participants to spontaneously use race to inform gaze following.

Second, and in contrast to attention cueing effects, incidental learning of trust from the predictive gaze patterns of ignored faces was influenced by race. That is, trust learning was larger and more robust for own race faces. As noted above, there is a wealth of previous literature that suggests we might see a difference in incidental learning processes between faces of different races (Dasgupta et al., 2000; Elfenbein & Ambady, 2002a, 2002b; Meissner & Brigham,

2001). However, that this learning occurred even without differences in attentional cueing suggests that this effect does not arise as a result of differences in sensitivity to gaze leading to different disruptions of processing fluency (cf. Strachan et al, 2016).

There are a number of potential explanations for this effect – that this result is driven by out-group homogeneity, as participants are less likely to individuate other-race members than own-race members; or that they more efficiently encode and store face identity for own-race than other-race members. However, as noted above, there is evidence that other race faces are trusted less than own race faces (Stanley et al., 2011). In Experiment 1, participants showed a non-significant bias in pre-experiment trustworthiness ratings to judge own-race faces as more trustworthy than other-race faces, even though faces were initially matched for trustworthiness when they were first selected. Although this was not significant, it is still possible that subtle differences in preconceptions about trustworthiness could have driven different strategies of learning for different identities. To investigate this, we report data from an earlier, independently run experiment that directly tests the role of trustworthiness in this incidental learning effect.

Experiment 2

Much research has investigated the physical features of a face that predict how trustworthy it is perceived to be (Oosterhof & Todorov, 2009; Todorov, 2008; Todorov, Baron, & Oosterhof, 2008; Todorov, Pakrashi, & Oosterhof, 2009). Physiognomic facial configurations such as wider jaws, lower brow ridges and other signals that resemble emotional expressions (Oosterhof & Todorov, 2009) are processed quickly and automatically. Reliable ratings of attributes such as

trust can be observed after only 100ms (Willis & Todorov, 2006), and the features used to make these decisions are consistent enough that they can be visualised and predicted using image-based analysis of ambient (i.e. not posed) images (Vernon, Sutherland, Young, & Hartley, 2014).

Experiment 2 was designed and run independent of Experiment 1, and aimed to address how these physiognomic features may affect trust learning. In the experiment used to collect these data we manipulated the baseline trustworthiness of the face (high/low trustworthiness) and tested trust learning (in a similar way to race in Experiment 1). Such a manipulation would create expectations (e.g. that trustworthy people will cooperate while untrustworthy people will deceive) and these expectations may interact with incidental learning of trust from eye-gaze behaviour – for example, given the expectation that trustworthy people are better social partners, participants may be more inclined to incidentally learn about their behaviour for reference in future interactions. If differences in social learning from own- and other-race faces are due to different levels of trust, we would expect that this independent experiment would have found the same profile of learning as Experiment 1 (that is, greater learning for trustworthy than untrustworthy faces).

Methods

Participants

Participants were 30 students from Bangor University (29 female, M_{age} 20). No participants were removed on the basis of RT filters (as detailed in Experiment 1). The study was given ethical approval by the Bangor University ethics committee. Details of participants' racial identity were not collected for these data.

Stimuli, Design and Procedure

The face stimuli were taken from the Karolinska Database of Emotional Faces(KDEF; Lundqvist, Flykt, & Öhman, 1998). All faces were female and selected based on ratings from Oosterhof and Todorov (2008). Sixteen faces were selected, eight of which were the faces rated highest for trust and eight of which were rated lowest for trust.

Some details of this experiment differed slightly from Experiment 1, due to the fact that they were run independent of each other. At the beginning of the trial, a fixation cross appeared for 1500ms followed by a directly gazing face for 1500ms. The face then changed gaze direction and remained for 500ms after which an object appeared to the left or right hand side of the face and disappeared as soon as a response was made or until 3000ms elapsed. When a response was made, the object disappeared and the face gazed directly ahead again for 2000ms. These timings are shown in Figure 1b. During trustworthiness ratings, the scale was labelled with '*Very untrustworthy*' and '*Very trustworthy*' (respectively). The faces subtended approximately 7.57˚ horizontally and 10.23˚ vertically from a distance of 60cm. Stimuli were displayed at a screen resolution of 800 × 600 pixels in the cueing phase and at 640 × 480 pixels in the rating phases. The experiment was displayed on a 19" Iiyama Vision-master CRT display. All stimuli were presented on a grey background. All other details were the same as those described in Experiment 1.

Data Analysis

All details of data analysis are identical to those outlined in Experiment 1, with the exception that in this experiment no participants were removed on the

basis of pre-processing filters, and face trustworthiness replaced race as a factor in all analyses.

Results

Gaze cueing

The RT results of Experiment 2 are shown in Figure 2 (lower row). A 2x2 repeated measures ANOVA of face trustworthiness and validity found a main effect of validity ($F(1,29) = 16.34$, $p < .001$, $\eta^2_P = 0.36$), but a non-significant effect of face trustworthiness ($F(1,29) = 3.66$, $p = 0.066$, $\eta^2_P = 0.11$). There was no evidence of enhanced gaze following for trustworthy faces, as there was no interaction of trust and validity ($F(1,29) = 0.10$, $p = 0.750$, $\eta^2_P = 0.00$).

The accuracy results (calculated as percent correct) of this experiment are shown in Table 1. A 2x2 ANOVA of face trustworthiness and validity found no effect of validity ($F(1,29) = 1.63$, $p = 0.211$, $\eta^2_P = 0.05$), or of trustworthiness ($F(1,29) = 1.15$, $p = 0.293$, $\eta^2_P = 0.04$), and no interaction ($F(1,29) = 0.76$, $p = 0.391$, $\eta^2_P = 0.03$).

Trustworthiness Ratings

The results of trustworthiness ratings at the beginning and end of this experiment are shown in Figure 3 (bottom row). A 2x2x2 ANOVA looking at time, validity and trustworthiness found no main effect of time ($F(1,29) = 0.00$, $p = 0.962$, $\eta^2_P = 0.00$), or validity ($F(1,29) = 4.09$, $p = 0.053$, $\eta^2_P = 0.12$), but did find a main effect of face trustworthiness on judgements ($F(1,29) = 123.59$, $p < .001$, $\eta^2_P = 0.81$). A significant interaction of time and validity was found ($F(1,29) = 11.11$, $p = 0.002$, $\eta^2_P = 0.28$), indicating that there was significant learning of trust over time as a function of gaze cueing behaviour. However, no other interactions were

significant, including the crucial three-way interaction of time, validity and trust ($F(1,29) = 0.01$, $p = 0.940$, $\eta^2_P = 0.00$; all other Fs<1).

We broke this down into separate 2-way ANOVAs that looked at the effects of validity and trustworthiness at the beginning and the end of the experiment, separately. At the beginning of the experiment, there was no main effect of validity ($F(1,29) = 1.29$, $p = 0.266$, $\eta^2_P = 0.04$), as participants had not been exposed to faces' valid or invalid behaviours at this point. There was, however, a large difference in pre-ratings of trust assigned to high- ($M = 21.28$, *s.d.* = 38.94) and low-trustworthiness faces ($M = -24.65$, *s.d.* = 37.30) and as expected this effect was significant ($F(1,29) = 122.01$, $p < .001$, $\eta^2_P = 0.81$) . There was no interaction of validity and trustworthiness ($F(1,29) = 0.01$, $p = 0.919$, $\eta^2_P = 0.00$).

At the end of the experiment there was a main effect of validity ($F(1,29) = 8.08$, $p = 0.008$, $\eta^2_P = 0.22$) due to incidental learning of patterns of gaze behaviour, and again a main effect of trustworthiness ($F(1,29) = 44.78$, $p < .001$, $\eta^2_P = 0.61$). However, importantly, in this experiment there was no significant interaction between the two ($F(1,29) = 0.00$, $p = 0.974$, $\eta^2_P = 0.00$) confirming that incidental learning of trust is equivalent for high and low trustworthy faces. At the end of the experiment there were significant differences between valid ($M = 28.24$, *s.d.* = 24.58) and invalid identities both with trustworthy faces ($M = 9.65$, *s.d.* = 32.86; $t(29) = 2.40$, 95%CI [2.77, 34.41], $p = .023$) and also those with untrustworthy faces (valid: $M = -12.86$, *s.d.* = 33.35; invalid: $M = -31.23$, *s.d.* = 31.11; $t(29) = 2.66$, 95%CI [4.25, 32.49], $p = .013$; Bonferroni-corrected α: .025).

The results of this experiment offer several points of interpretation. First, shifts of attention caused by gaze cues are not affected by the trustworthiness of

the face. These gaze cueing effects are similar to those of Experiment 1 where no differences in attention cueing were observed between own and other race faces.[2] Second, although trustworthiness judgements are heavily driven by the physical appearance of the face, in line with previous research (Sutherland et al., 2013; Todorov, 2008; Todorov et al., 2008; Vernon et al., 2014), this has no effect on the incidental learning of trust from eye-gaze behaviour.

Therefore, we can conclude that the contrast in incidental learning between own and other race faces observed in Experiment 1 is not determined by differences in levels of trustworthiness. In Experiment 2 a direct manipulation of trust based on physiognomic properties did not detect any effects on trust learning from gaze behaviour when viewing only Caucasian faces. Therefore, the hypothesis that trust learning is reduced in other race faces compared with own race faces due to differences in participants' initial feelings of trust is not supported.

General Discussion

The current study reports the results of two experiments exploring how the identity of a cueing face – and the higher order social information that this carries – can affect orienting of attention and the incidental learning of trust from gaze cues. In both experiments we observe that cueing of attention to the right and left by eye-gaze is unaffected by the nature of the face, whether race or trustworthiness.

This supports previous evidence that gaze cues orient attention in a very fast and automatic manner that is difficult to inhibit (Driver et al., 1999; Freebody & Kuhn, 2016; Frischen & Tipper, 2004). Although previous research has found that cueing can be mediated by factors such as social status (Dalmaso,

Pavan, Castelli, & Galfano, 2012), dominance (Jones et al., 2010), familiarity (Deaner, Shepherd, & Platt, 2007) and, indeed, race (Dalmaso et al., 2015; Pavan et al., 2011) and trustworthiness (Petrican et al., 2013; Süßenbach & Schönbrodt, 2014), we found no evidence that participants spontaneously considered either of the latter features when processing gaze, suggesting that these mediating effects may rely on the context (e.g. perceived threat) in which participants find themselves experiencing gaze cues.

This is particularly striking in Experiment 2, where we failed to replicate previous research that shows that trustworthiness can affect gaze cueing (Petrican et al., 2013; Süßenbach & Schönbrodt, 2014). There may be a variety of reasons for these contrasts: with regards to Petrican et al., we recruited young adults for all experiments reported here, where they found that trustworthiness affected gaze cueing only in older adults. While Süßenbach and Schönbrodt also used younger adults, they used affectively valenced target stimuli where both of the current experiments used neutral household items, a simpler left/right discrimination task where ours was a more demanding category identification judgement, and (perhaps most importantly) used familiar faces that were known from background information to be trustworthy or untrustworthy (characters in films), whereas ours used unknown faces that differed in perceptual physiognomic features. Further research would be needed to identify which of these design features contributes to whether or not trustworthiness affects the magnitude of gaze cueing effects, but our present findings certainly suggest that if people are susceptible to trustworthiness information during gaze cueing, they do not invariably use it spontaneously whenever it is available.

However, the main focus of our study was to examine how the nature of the face, whether own- or other-race (or high- or low-trust), would influence the learning of trust from gaze behaviour. We predicted that trust learning would be less efficient when viewing other race faces. This was based on the idea that during the task where faces are irrelevant and to-be-ignored, an association has to be learned between a specific face identity and the pattern of eye-gaze it produces, and that this would be processed more efficiently for own- than other-race faces. Although no differences were found in susceptibility to gaze cues, we nonetheless found that race affected how participants learned about the trustworthiness of individuals. This suggests that these processes use different underlying mechanisms – a fast, attention-orienting mechanism that processes gaze cues and does not spontaneously take the race of the face into account, and another mechanism that reviews gaze behaviour and incorporates this into a stable representation of that particular identity for use in future social decisions.

It follows that the association between identity and gaze behaviour will be more easily learned if there is a strong/specific representation of the face identity. Strachan and Tipper (2017) confirmed this by manipulating the strength of face identity representations, demonstrating that stronger representations resulted in greater learning of trust from gaze behaviour. There is extensive prior research demonstrating that other race faces are identified and remembered less efficiently than own race faces (see Meissner & Brigham, 2001, for a review). As such, this could be a plausible explanation that future research may look to investigate further.

It would also be interesting for future research to investigate how participants may use racial group membership differently depending on their

own identity. In Experiment 1 we used exclusively Caucasian participants. The reason we did not include East Asian participants as a contrast was because within the sample population (undergraduate students at the University of York), East Asian participants are a minority group, and there is some evidence that people process group dynamics differently on the basis of whether their in-group is a majority or minority (Elfenbein & Ambady, 2002b). However, it would be interesting to explore in future research whether the status of participants (as members of majority or minority groups) influenced participants' sensitivity to gaze behaviour in an incidental learning scenario. Such future research may wish to contrast social learning in such a minority population with a matched majority (e.g. Chinese participants living in the UK compared with Chinese participants living in China).

With the inclusion of Experiment 2, which manipulated face trustworthiness, we were able to examine the role of trust in such learning processes. It was noted that previous work has reported less trust of other race individuals (e.g., Stanley et al., 2011), although there were only trends for this pattern in the current study. Although we hypothesised that these subtle differences in preconceptions about trust between racial groups could still have played a role in the different learning profiles seen in Experiment 1, analysis of this independent dataset demonstrated that initial trust of a face was not influential. That is, learning of trust from gaze was not impaired in low-trust faces, suggesting that the results of Experiment 1 cannot be explained by different levels of trust associated to in-group and out-group members.

Consequently, our findings confirm that while the fixed physiognomic properties of a face are a strong predictor of trustworthiness judgements, and

while incidental learning of cueing contingencies has an effect, it does not override or interact with this more salient perceptual information, making it unlikely that this feature can explain the results of Experiment 1. Rather, our gaze manipulation moderates initial trust ratings in similar ways. However, note that the gaze learning is incidental while faces are ignored and we are examining effects on judgements at a later time where there are no visible cues to prior deception. This is in sharp contrast to other more in-the-moment manipulations of trust such as face emotion (which does affect trust learning; Bayliss et al., 2009; Strachan et al., 2016), which are salient physical properties of a face that are present while trust judgements are actually made.

In conclusion, our studies demonstrated a number of features of incidental learning of trust from gaze cues. Learning is incidental in that participants are ignoring the faces, and hence these demanding learning conditions are influenced by the robustness of the representations that have to be associated. In Experiment 1, learning is impaired for other race faces that have weaker representations of each identity. In contrast, Experiment 2 demonstrated that the race effects are probably not driven by initial trustworthiness of own versus other races. When the faces are all Caucasian, large differences in trust do not influence incidental learning from gaze behaviour.

Footnotes

1. The primary manipulation in this experiment was that of racial group, but our stimuli also included faces that varied according to gender. This also gave an additional orthogonal group membership factor of face gender that could have affected results. Previous research has found no evidence that gender affects subsequent trust learning (Manssuer et al., 2016), but it was possible that in this paradigm the presence of an additional group dimension (race) also made gender a salient distinction. A 2x2x2 repeated measures ANOVA with time, validity, and gender in place of race as fixed factors found no main effect of gender ($F(1,29) = 2.54$, $p = 0.122$, $\eta^2_P = 0.08$) and no interactions of gender with either time ($F(1,29) = 0.02$, $p = 0.880$, $\eta^2_P = 0.00$) or validity ($F(1,29) = 1.51$, $p = 0.229$, $\eta^2_P = 0.05$) and no three-way interaction ($F(1,29) = 0.01$, $p = 0.935$, $\eta^2_P = 0.00$). The same held true when examining only female participants (26/30): there was no main effect of gender ($F(1,25) = 1.50$, $p = 0.232$, $\eta^2_P = 0.06$) and no interactions of gender with either time ($F(1,25) = 0.01$, $p = 0.918$, $\eta^2_P = 0.00$) or validity ($F(1,25) = 2.68$, $p = 0.114$, $\eta^2_P = 0.10$) and no three-way interaction ($F(1,25) = 0.04$, $p = 0.849$, $\eta^2_P = 0.00$).
2. Gaze cueing effects looked largely similar across both experiments. A 2x2 mixed ANOVA on collapsed data from these experiments, with experiment (1/2) as a between-subjects factor and validity (valid/invalid) as a within-subjects factor found a main effect of validity ($F(1,58) = 34.03$, $p < .001$, $\eta^2_P = 0.37$), and a main effect of experiment ($F(1,58) = 7.17$, $p = 0.010$, $\eta^2_P = 0.11$). This effect was driven by the fact that RTs were longer overall in one experiment than the other. However, this effect

did not interact with gaze cueing across the different experiments ($F(1,58) = 0.30$, $p = 0.589$, $\eta^2_P = 0.01$), meaning that sensitivity to gaze cues did not differ significantly across the two experiments.

Acknowledgements

This work was supported by the Economic and Social Research Council [ES/000012/1].

References

Allen, V. L., & Wilder, D. A. (1975). Categorization, belief similarity, and intergroup discrimination. *Journal of Personality and Social Psychology*, *32*(6), 971–977. https://doi.org/10.1037/0022-3514.32.6.971

Bayliss, A. P., Griffiths, D., & Tipper, S. P. (2009). Predictive gaze cues affect face evaluations: The effect of facial emotion. *European Journal of Cognitive Psychology*, *21*(7), 1072–1084. https://doi.org/10.1080/09541440802553490

Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments: Should eye trust you? *Psychological Science*, *17*(6), 514–520. https://doi.org/10.1111/j.1467-9280.2006.01737.x

Caldara, R., Rossion, B., Bovet, P., & Hauert, C.-A. (2004). Event-related potentials and time course of the "other-race" face classification advantage. *Neuroreport*, *15*(5), 905–10. https://doi.org/10.1097/00001756-200404090-00034

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*(5), e97644. https://doi.org/10.1371/journal.pone.0097644

Chen, Y., & Zhao, Y. (2015). Intergroup threat gates social attention in humans. *Biology Letters*, *11*(2), 20141055. https://doi.org/10.1098/rsbl.2014.1055

Chen, Y., Zhao, Y., Song, H., Guan, L., Wu, X., Bayliss, A. P., … Jenkinson, M. (2017). The neural basis of intergroup threat effect on social attention. *Scientific Reports*, *7*, 41062. https://doi.org/10.1038/srep41062

Dalmaso, M., Galfano, G., & Castelli, L. (2015). The Impact of Same- and Other-

Race Gaze Distractors on the Control of Saccadic Eye Movements. *Perception*, *44*(8–9), 1020–1028. https://doi.org/10.1177/0301006615594936

Dalmaso, M., Pavan, G., Castelli, L., & Galfano, G. (2012). Social status gates social attention in humans. *Biology Letters*, *8*(3), 450–452.

Dasgupta, N., Mcghee, D. E., Greenwald, A. G., & Banaji, M. R. (2000). Automatic Preference for White Americans: Eliminating the Familiarity Explanation. *Journal of Experimental Social Psychology*, *328*(3), 316–328. https://doi.org/10.1006/jesp.1999.1418

Deaner, R. O., Shepherd, S. V, & Platt, M. L. (2007). Familiarity accentuates gaze cuing in women but not men. *Biology Letters*, *3*(1), 64–67. https://doi.org/10.1098/rsbl.2006.0564

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, *6*(5), 509–540. https://doi.org/10.1080/135062899394920

Elfenbein, H. A., & Ambady, N. (2002a). Is there an in-group advantage in emotion recognition? *Psychological Bulletin*, *128*(2), 243–249. https://doi.org/10.1037/0033-2909.128.2.243

Elfenbein, H. A., & Ambady, N. (2002b). On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological Bulletin*, *128*(2), 203–235. https://doi.org/10.1037/0033-2909.128.2.203

Freebody, S., & Kuhn, G. (2016). Own-age biases in adults' and children's joint attention: Biased face prioritization, but not gaze following! *The Quarterly Journal of Experimental Psychology*, 1–9. https://doi.org/10.1080/17470218.2016.1247899

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*(3), 490–495. https://doi.org/10.3758/BF03208827

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694–724. https://doi.org/10.1037/0033-2909.133.4.694

Frischen, A., & Tipper, S. P. (2004). Orienting attention via observed gaze shift evokes longer term inhibitory effects: implications for social interactions, attention, and memory. *Journal of Experimental Psychology. General*, *133*(4), 516–33. https://doi.org/10.1037/0096-3445.133.4.516

Galfano, G., Dalmaso, M., Marzoli, D., Pavan, G., Coricelli, C., & Castelli, L. (2012). Eye gaze cannot be ignored (but neither can arrows). *The Quarterly Journal of Experimental Psychology*, *65*(10), 1895–1910. https://doi.org/10.1080/17470218.2012.663765

Jones, B. C., DeBruine, L. M., Main, J. C., Little, A. C., Welling, L. L. M., Feinberg, D. R., & Tiddeman, B. P. (2010). Facial cues of dominance modulate the short-term gaze-cuing effect in human observers. *Proceedings. Biological Sciences / The Royal Society*, *277*(1681), 617–624. https://doi.org/10.1098/rspb.2009.1575

Lickel, B., Hamilton, D. L., & Sherman, S. J. (2001). Elements of a lay theory of groups: Types of groups, relational styles, and the perception of group entitativity. *Personality and Social Psychology Review*, *5*(2), 129–140. https://doi.org/10.1207/S15327957PSPR0502_4

Lin, M. H., Kwan, V. S. Y., Cheung, A., & Fiske, S. T. (2005). Stereotype Content Model Explains Prejudice for an Envied Outgroup: Scale of Anti-Asian

American Stereotypes. *Personality and Social Psychology Bulletin*, *31*(1), 34–47. https://doi.org/10.1177/0146167204271320

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet*, 91–630.

Manssuer, L. R., Pawling, R., Hayes, A. E., & Tipper, S. P. (2016). The role of emotion in learning trustworthiness from eye-gaze: Evidence from facial electromyography. *Cognitive Neuroscience*, *7*(1–4), 82–102. https://doi.org/10.1080/17588928.2015.1085374

Manssuer, L. R., Roberts, M. V, & Tipper, S. P. (2015). The late positive potential indexes a role for emotion during learning of trust from eye-gaze cues. *Social Neuroscience*, *10*(6), 635–650. https://doi.org/10.1080/17470919.2015.1017114

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychology, Public Policy, and Law*, *7*(1), 3–35. https://doi.org/10.1037/1076-8971.7.1.3

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), 11087–92. https://doi.org/10.1073/pnas.0805664105

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

Pavan, G., Dalmaso, M., Galfano, G., & Castelli, L. (2011). Racial group membership is associated to gaze-mediated orienting in Italy. *PLoS ONE*, *6*(10), e25608. https://doi.org/10.1371/journal.pone.0025608

Petrican, R., English, T., Gross, J. J., Grady, C., Hai, T., & Moscovitch, M. (2013). Friend or foe? Age moderates time-course specific responsiveness to trustworthiness cues. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *68*(2), 215–223. https://doi.org/10.1093/geronb/gbs064

Rogers, R. D., Bayliss, A. P., Szepietowska, A., Dale, L., Reeder, L., Pizzamiglio, G., … Tipper, S. P. (2014). I want to help you, but I am not sure why: gaze-cuing induces altruistic giving. *Journal of Experimental Psychology. General*, *143*(2), 763–77. https://doi.org/10.1037/a0033677

Sessa, P., & Dalmaso, M. (2016). Race perception and gaze direction differently impair visual working memory for faces: An event-related potential study. *Social Neuroscience*, *11*(1), 91–107. https://doi.org/10.1080/17470919.2015.1040556

Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(19), 7710–5. https://doi.org/10.1073/pnas.1014345108

Strachan, J. W. A., Kirkham, A. J., Manssuer, L. R., & Tipper, S. P. (2016). Incidental learning of trust: Examining the role of emotion and visuomotor fluency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(11), 1759–1773. https://doi.org/10.1037/xlm0000270

Strachan, J. W. A., & Tipper, S. P. (2017). Examining the durability of incidentally learned trust from gaze cues. *The Quarterly Journal of Experimental Psychology*, *70*(10), 2060–2075. https://doi.org/10.1080/17470218.2016.1220609

Strohminger, N., Gray, K., Chituc, V., Heffner, J., Schein, C., & Heagins, T. B. (2015). The MR2: A multi-racial, mega-resolution database of facial stimuli. *Behavior Research Methods*. https://doi.org/10.3758/s13428-015-0641-9

Süßenbach, F., & Schönbrodt, F. (2014). Not afraid to trust you: Trustworthiness moderates gaze cueing but not in highly anxious participants. *Journal of Cognitive Psychology*, *26*(6), 670–678. https://doi.org/10.1080/20445911.2014.945457

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*(1), 105–118. https://doi.org/10.1016/j.cognition.2012.12.001

Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2016). Facial first impressions from another angle: How social judgements are influenced by changeable and invariant facial properties. *British Journal of Psychology*. https://doi.org/10.1111/bjop.12206

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149–178. https://doi.org/10.1002/ejsp.2420010202

Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*, *1124*, 208–224. https://doi.org/10.1196/annals.1440.012

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social Cognitive and Affective Neuroscience*, *3*(2), 119–127. https://doi.org/10.1093/scan/nsn009

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating Faces on Trustworthiness After Minimal Time Exposure. *Social Cognition*, *27*(6), 813–833. https://doi.org/10.1521/soco.2009.27.6.813

Vernon, R. J. W., Sutherland, C. A. M., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences*, *111*(32), E3353-3361. https://doi.org/10.1073/pnas.1409860111

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592–598. https://doi.org/10.1111/j.1467-9280.2006.01750.x

Table 1. Accuracy rates (percent correct with *standard error*) averaged across subjects in Experiment 1 (own race/other race faces) and Experiment 2 (high/low trustworthy faces) for valid and invalid trials.

		Valid	*Invalid*
Experiment 1	*Own race*	96.90 *(3.17)*	96.55 *(3.33)*
	Other race	96.42 *(3.40)*	96.50 *(3.36)*
Experiment 2	*High trust*	96.33 *(3.43)*	96.25 *(3.47)*
	Low trust	97.50 *(2.85)*	96.25 *(3.47)*

Figure captions

Figure 1. a. Examples of the four different conditions in which faces were presented in Experiment 1: out-group valid, out-group invalid, in-group valid and in-group invalid. In Experiment 2, the conditions were faces that were previously rated as high and low in trustworthiness. b. Schematic of two gaze cueing trials; one out-group valid (top row) and one in-group invalid (bottom row). The duration of each trial event is displayed along the bottom for Experiment 1 and Experiment 2. If participants made a mistake an error tone would play between the last two trial events.

Figure 2. Reaction time in milliseconds to valid (light grey) and invalid (dark grey) trials in Experiment 1 (top plot; own race trials on the left, other race trials on the right) and Experiment 2 (bottom plot; highly trustworthy faces on the left, low trustworthy faces on the right). Error bars show ±1 within-subjects standard error.

Figure 3. Trustworthiness ratings from Experiment 1 (top row) with own race (left plot) and other race faces (right plot), and Experiment 2 (bottom row) with faces high in trustworthiness (left plot) and low in trustworthiness (right plot). Ratings are shown over time separately for valid (dotted lines) and invalid (solid lines) trials. Error bars show ±1 within-subjects standard error.