Date of acceptance: 5th March 2021.

Text

The interplay between gaze cueing and facial trait impressions

Journal:	*Quarterly Journal of Experimental Psychology*
Manuscript ID	QJE-STD-20-255.R2
Manuscript Type:	Standard Article
Date Submitted by the Author:	01-Mar-2021
Complete List of Authors:	Mattavelli, Giulia; Scuola Universitaria Superiore Pavia, Humanities and Life Sciences Romano, Daniele; University of Milan at Bicocca Young, Andrew; University of York, Ricciardelli, Paola; Università di Milano Bicocca, Dipartimento di Psicologia
Keywords:	Gaze cueing effect, social traits, trustworthiness, dominance, incidental learning

DOI: 10.1177/17470218211007791

The interplay between gaze cueing and facial trait impressions

Giulia Mattavelli[1*], Daniele Romano[2,3], Andrew W Young[4], Paola Ricciardelli[2,3]

[1] *Scuola universitaria superiore IUSS, Pavia, Italy*

[2] *Dipartimento di Psicologia, Università degli studi di Milano-Bicocca, Milano, Italy*

[3] *NeuroMi (Neuroscience Center), University of Milano-Bicocca, Milan, Italy*

[4] *Department of Psychology, University of York, York, UK*

* corresponding author:

Giulia Mattavelli

email: giulia.mattavelli@iusspavia.it

Address: Palazzo del Broletto, Piazza della Vittoria 15, 27100 Pavia

DOI: 10.1177/17470218211007791

Abstract

The gaze cueing effect involves the rapid orientation of attention to follow the gaze direction of another person. Previous studies reported reciprocal influences between social variables and the gaze cueing effect, with modulation of gaze cueing by social features of face stimuli and modulation of the observer's social judgments from the validity of the gaze cues themselves. However, it remains unclear which social dimensions can affect - and be affected by - gaze cues. We used computer-averaged prototype face-like images with high and low levels of perceived trustworthiness and dominance to investigate the impact of these two fundamental social impression dimensions on the gaze cueing effect. Moreover, by varying the proportions of valid and invalid gaze cues across three experiments, we assessed whether gaze cueing influences observers' impressions of dominance and trustworthiness through incidental learning. Bayesian statistical analyses provided clear evidence that the gaze cueing effect was not modulated by facial social trait impressions (Experiments 1-3). On the other hand, there was uncertain evidence of incidental learning of social evaluations following the gaze cueing task. A decrease in perceived trustworthiness for non-cooperative low dominance faces (Experiment 2) and an increase in dominance ratings for faces whose gaze behaviour contradicted expectations (Experiment 3) appeared, but further research is needed to clarify these effects. Thus, this study confirms that attentional shifts triggered by gaze direction involve a robust and relatively automatic process, which could nonetheless influence social impressions depending on perceived traits and the gaze behaviour of faces providing the cues.

Keywords: Gaze cueing effect, social traits, trustworthiness, dominance, incidental learning.

DOI: 10.1177/17470218211007791

Introduction

Eye gaze is a crucial source of information in social interactions; it provides cues about the intentions and emotions of others, as well as signalling where the other person is attending and hence the presence of relevant objects in the environment (Mundy & Newell, 2017; Calder et al., 2002). In particular, the ability to follow others' gaze, a phenomenon known as joint attention, develops in early infancy and is a fundamental aspect of social cognition (Mundy & Newell, 2017; Tomasello et al., 2005).

In adults, joint attention has been extensively investigated by means of the gaze cueing paradigm, a task in which participants are asked to detect or discriminate stimuli presented to the right or left side of a central face with the eyes gazing to the same (i.e. congruent or valid cue) or opposite (i.e. incongruent or invalid cue) location as the target (Friesen & Kingstone, 1998; Driver et al., 1999). A consistent finding has involved faster responses in trials with valid gaze cues, demonstrating that eye gaze automatically triggers an attentional shift of the observer in the same direction (Frieschen et al., 2007).

In the past decade research on the gaze cueing effect has moved beyond establishing the basic effect to investigate the impact of different manipulations concerning the features of the face stimuli or the contextual relationship between observers and stimuli (e.g. priming conditions or person descriptions) (see Dalmaso et al., 2020a for review). The central question of these studies has thus become whether the gaze cueing effect depends on automatic processes triggered independently from social variables, or whether it can be influenced by top-down mechanisms which might affect the attentional shift depending on the perception and processing of social properties. As reported in a recent review (Dalmaso et al., 2020a), different studies have demonstrated that the gaze cueing effect can sometimes be influenced by contextual information concerning the faces (e.g. social status, Dalmaso et al, 2012; Ciardo et al., 2021) and face properties such as apparent age, ethnicity, emotional expressions and social traits, supporting the hypothesis of top-down modulation. However, results are not always consistent and suggest potentially complex interactions between context, stimuli and observers (Ciardo et al., 2014; Hietanen & Leppänen, 2003; Ohlsen et al., 2013; Pavan et al., 2011; Pecchinenda et al., 2008; Bonifacci et al., 2008). In particular, concerning social trait impressions, modulations of the gaze cueing effect have been investigated in relation to dominance and trustworthiness. These are considered two main trait dimensions in the facial impressions literature (Oosterhof & Todorov, 2008). Indeed, studies of trait judgments using computer-generated faces have

DOI: 10.1177/17470218211007791

identified trustworthiness and dominance as the two best fitting dimensions in a widely-used model resulting from a data-driven statistical approach, leading to the suggestion that face evaluation is related to adaptive mechanisms developed to characterise others in terms of their approachability (i.e. trustworthiness) and strength and weakness (i.e. dominance) (Oosterhof & Todorov, 2008; Todorov et al., 2008). This dimensional model has been replicated in subsequent studies, although a third dimension related to youthful-attractiveness and differences between male and female faces have also emerged in studies asking for evaluations of natural and more variable face photographs (Oh et al., 2019; Sutherland et al., 2013; 2015; Vernon et al., 2014). Importantly, studies have shown that impressions of trustworthiness and dominance are rapidly formed, with consistent evaluations of stimuli presented even as briefly as 100 ms (Ballew & Todorov, 2007; Bar, Neta, & Linz, 2006; Olivola & Todorov, 2010; South Palomares & Young, 2018; Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006) suggesting that, like the gaze cueing effect itself, impressions of faces are relatively automatically generated.

A larger gaze cueing effect for more dominant faces has been reported in two studies, which manipulated facial cues of dominance in prototype faces or presented faces previously rated as extremely high or low dominant (Jones et al., 2010; Ohlsen et al., 2013). However, these studies manipulated masculinity-femininity of faces or presented male-dominant and female-non-dominant faces, thus varying dominance in combination with variation in gender. A different study showed that faces eliciting greater gaze cueing effects were rated as more dominant by an independent group of observers, supporting the hypothesis that perceived dominance modulates responses in gaze cueing tasks (Jones et al., 2011). On the other hand, studies assessing the impact of face trustworthiness in gaze cueing paradigms have provided inconsistent data. Larger effects for trustworthy faces were reported in two studies which used computer-generated faces from Oosterhof and Todorov's (2008) set (Petrican et al., 2013) or manipulated the perceived trustworthiness of unfamiliar faces by providing personality descriptions of the characters (Süßenbach & Schönbrodt, 2014). However, these findings were not replicated in other studies which manipulated trustworthiness of stimuli by means of descriptive vignettes, selecting faces with high or low rated trustworthiness, or fixing the probability of valid and invalid cues associated to different faces so that some stimuli always provided valid cues and others invalid cues (Bayliss & Tipper, 2006; King et al., 2011; Strachan et al., 2017). Indeed, a recent review on the effect of trust on gaze cueing of attention highlighted the contradictory results in the

DOI: 10.1177/17470218211007791

previous literature, pointing to methodological differences between studies related to features of the stimuli, trust manipulation and sample characteristics. The review encouraged further research to unveil the relationship between social judgment and orienting of attention (Barbato et al., 2020). It is also important to consider the temporal dynamics of gaze cueing and of the possible impact of social variables. This facet has been investigated by manipulating the stimulus-onset asynchrony (SOA), namely the temporal interval between the cueing phase (i.e. the lateral orientation of eye gaze) and the onset of the lateralised target. The majority of studies assessed the impact of SOA in a time renge between about 100 and 1200 ms. The gaze cueing effect has been reported with SOAs as short as 105 ms, suggesting that orientation of attention is reflexive (Friesen and Kingstone, 1998; Driver et al., 1999). Increasing SOA produced an overall reduction of reaction times, likely related to a foreperiod effect, whereas the gaze cue mediated facilitation disappeared with a longer SOA of 1200 ms (Friesen et al., 2004; Frischen et al., 2007). However, the is also evidence that in certain conditions gaze cueing effect can be detected even at the very short SOA of 14 ms or following 3 minutes SOA (Hietanen and Leppänen, 2003; Frischen & Tipper, 2006). Evidence on any modulatory influence of social variables is inconsistent, as previous studies reported either effects of face dominance at short SOA rather than long SOA (i.e. 200 ms vs 400 ms; Jones et al., 2010) or effects of face trustworthiness at long SOA rather than short SOA (i.e. 600 ms vs 100 ms; Petrican et al., 2013). Therefore, it is a matter of debate whether social modulation might occur in a reflexive manner, or through top-down control requiring a longer SOA (Dalmaso et al., 2020a; Barbato et al., 2020).

Another relevant aspect to acknowledge in the relationship between joint attention and social variables is that eye gaze plays a crucial role in social evaluation. Indeed, several studies have shown that gaze direction affects facial expression processing and social judgments (Mason et al., 2005; Milders et al., 2011; Sander et al., 2007). For example, faces with gaze directed to the perceiver are rated as more dominant (Main et al., 2009), more attractive and more trustworthy (Ewing et al., 2010; Kaisler & Leder, 2016) than faces with averted eyes. Therefore, if we consider gaze cueing paradigms as mimicking some basic properties of social interaction, albeit in a constrained experimental setting, it is interesting to investigate the influence of gaze cues in two different directions: on the one hand how social attributes may modulate the gaze cueing effect, as well as on the other hand whether gaze cues have an impact on social judgments. In this respect it is noteworthy that implicit learning related to gaze cueing has been demonstrated in previous studies showing

DOI: 10.1177/17470218211007791

that perceived face trustworthiness increased for faces which consistently provided valid cues, whereas trustworthiness decreased for faces which always deceived (Bayliss & Tipper 2006; Bayliss et al., 2009; Strachan et al., 2017; Manssuer et al., 2015; 2016). However, it remains unclear which face properties can affect implicit learning and produce changes in social evaluations following gaze cueing paradigms. Trust learning has been found to be affected by ethnicity of stimuli and facial expressions, but not by baseline level of trustworthiness (Bayliss et al., 2009; Strachan et al., 2017; Manssuer et al., 2015; 2016). Crucially, only trustworthiness judgments have been investigated in previous studies, whereas incidental learning of other social trait impressions from eye gaze cues has not been explored so far.

Although much of the research literature treats trait impressions as intrinsic properties of the faces themselves, an important recent finding has been that impressions of unfamiliar faces are actually highly image-dependent; for example, the same face can look trustworthy in one image and untrustworthy in another image (Jenkins, White, Van Montfort & Burton, 2011; Todorov & Porter, 2014). This shows that although facial impressions can be consensual, in the sense that different observers will show substantial agreement, they cannot reflect valid inferences about stable personality traits. Instead, it seems more likely that facial impressions derive from situation-based appraisals of a person's behaviour (e.g., a person can be much more dominant in some situations compared to others) which are then misinterpreted as cues to an underlying stable personality. This misinterpretation forms a parallel with the 'fundamental attribution error' of interpreting what are actually situationally-driven behaviours as if they were personality dispositions, as noted in many studies in social psychology (Todorov, 2017).

The differences in impressions created by different facial images are known to reflect the interaction of multiple cues (Vernon et al., 2014; Young, 2018); for example, a smile can make a face look warm and trustworthy when combined with one set of cues and yet shifty and untrustworthy when combined with different cues. A powerful way to exploit this property is to use computer manipulation techniques to create an averaged representation from multiple different images rated as high or low on a given trait, to arrive at a prototypical face-like image that carries as many as possible of the covarying cues that create the stereotypical impression (Sutherland et al., 2013; Sutherland et al., 2015; Sutherland, Rhodes & Young, 2017). Here, we made use of such face-like prototype images derived from previous research by Sutherland et al. (2015). The usefulness of this image-averaging technique has already been demonstrated in behavioural

DOI: 10.1177/17470218211007791

(Sutherland et al., 2013, 2015), neuropsychological (Sprengelmeyer et al., 2016), and functional brain imaging studies (Mattavelli et al., 2012).

The aim of this study was therefore two-fold: i) to test whether the gaze cueing effect is modulated by the social trait impressions created by the face images providing the cues and ii) to test whether gaze cues can change the attribution of social trait impressions of faces depending on the validity of the cues provided. To do so, we used computer-averaged prototype face-like images created to represent high or low trustworthy male and female individuals and high or low dominant male and female individuals (Sutherland et al., 2015) in three experiments. Each experiment began with an initial phase involving a rating paradigm to establish impressions of trustworthiness and dominance created by the face-like images. This was followed by a second phase using a gaze cueing paradigm that manipulated the proportion of valid or invalid cues provided by each type of image. Finally, the images were again rated on trustworthiness and dominance scales to assess the effect of the gaze cueing contingency on subsequent social trait evaluation.

Methods

Participants

A total of 121 university students participated in this study, in three different experiments: 40 participants in Experiment 1 (19 male, 21 females, mean age =23.03 ± 1.58), 40 participants in Experiment 2 (21 males, 19 females, mean age = 24.47 ± 4.33) and 41 participants in Experiment 3 (21 males, 20 females, mean age = 22.27 ± 2.85). The sample size was determined according to previous studies that investigated social modulation of the gaze cueing effect and incidental learning with samples of about 30 participants (Strachan et al., 2017; Pecchinenda et al., 2008; Ciardo et al., 2014; Pavan et al., 2011; Jones et al., 2010). We acknowledged that determining the sample size according to the typical numerosity adopted in the literature is not the best way to define the sample size, thus we performed a sequential analysis, which evaluate the evidence favouring the different models at the increasing of the sample size. This allowed to assess the reliability and the potential limits of our results (see online Supplementary material for detailed methodology and results of sequential analysis).

DOI: 10.1177/17470218211007791

The study was conducted in accordance with the ethical standards laid down in the 2013 Declaration of Helsinki and Oviedo convention. The study was also approved by the Ethics committee of the University of Milano-Bicocca, and all participants gave informed consent prior to participation.

Stimuli and procedure

Stimuli consisted of face-like images (for convenience referred to as 'faces' henceforth) taken from the study by Sutherland et al. (2015). These were created from a large set of everyday face photographs with a computer-averaging procedure to generate stereotypical prototype representations of male and female individuals perceived as high and low in trustworthiness (HT and LT) and dominance (HD and LD). Full details of the procedure used to create these stimuli are given by Sutherland et al. (2013, 2015). The eight original images (see Figure 1), which all had approximately central gaze direction, were modified with Adobe Photoshop software to create two new versions of each face with eye gaze directed to the right or left. Thus, we obtained 24 face stimuli, which were used in three different experiments that involved a gaze cueing paradigm preceded and followed by stimuli rating on trustworthiness and dominance dimensions.

In the gaze cueing paradigm participants were asked to discriminate the target letters L or T presented laterally to faces with eye gaze directed to the right or left congruently or incongruently with respect to the side of target presentation. Each trial started with a fixation cross at the centre of the screen for 500 ms, then one of the eight face stimuli with direct gaze appeared for 400 ms, followed by the same face with right or left averted gaze (cueing phase) for either 200 ms or 700 ms, after which the target letter appeared (see Figure 1). These SOAs of 200 and 700 ms for the gaze cueing were selected to test whether modulations of social trait impressions would occur at the shortest SOA in a reflexive manner, or through top-down control requiring a longer SOA (Jones et al., 2010; Dalmaso et al., 2020a). The two SOAs and the presentation of the L or T target to the right or left of the face were randomised and equally probable. A total of 1280 trials were presented in four blocks. The participants responded with index and middle fingers of their dominant hand pressing the up-arrow and the down-arrow, which corresponded to L or T targets in balanced order between participants. The proportion of valid and invalid cues provided by the gaze direction of each face in the cueing phase was manipulated across three experiments. In Experiment 1, each of the eight faces provided valid or invalid cues in 50% of trials in random order, such that all faces had an equal probability of being

DOI: 10.1177/17470218211007791

presented in the cueing phase with the eye gaze directed to the right or left, and congruently or incongruently with respect to the side of target presentation. In Experiment 2, HT and HD faces provided valid cues in 70% of trials and invalid cues in 30% of trials, whereas LT and LD faces provided valid cues in 30% of trials and invalid cues in 70% of trials. Then in Experiment 3, HT and HD faces were associated with valid cues in 30% of trials and with invalid cues in 70% of trials, whereas LT and LD faces were associated with valid cued in 70% of trials and with invalid cues in 30% of trials.

In the rating tasks, participants were asked to evaluate on 7-point Likert scales the 24 face stimuli (i.e. male and female HT, HD, LT, LD faces with direct, right-side and left-side gaze) for the trustworthiness and the dominance dimensions. Each face was presented at the centre of the screen with the scale from 1 (not trustworthy/dominant at all) to 7 (very trustworthy/dominant) below it, and the participants responded by pressing buttons from 1 to 7 on the keyboard. Ratings of trustworthiness and dominance were collected in two blocks counterbalanced across the participants and the order of the stimuli was randomised within each block.

During the experiments, the participants were seated at a desk in a dimly lit room with their head stabilised on a chin-rest in front of a computer screen. Tasks were presented with E-prime software (V.2, Psychology Software Tools), which recorded accuracy and latencies of responses. The procedure was the same for the three experiments: the participants first completed the rating task, then completed the gaze cueing paradigm, followed by the second presentation of the rating task. The first rating allowed us to check the effect of manipulating the gaze direction in the stimuli, whereas the second rating allowed us to test the impact of incidental learning on social impressions of the faces following the gaze cueing paradigm.

[Figure 1 about here]

Statistical analyses

A Bayesian statistical approach was used for data analysis (Wagenmakers et al., 2018). Based on our two main objectives we wanted to test the hypotheses that (i) gaze cueing was influenced by face trustworthiness and dominance at different SOAs and that (ii) the perceived trustworthiness and dominance of faces was influenced by incidental learning from performing the gaze cueing task. To assess evidence for and against

DOI: 10.1177/17470218211007791

the presence of these effects we ran Bayesian repeated measures ANOVAs using JASP software (Version 0.9.2) (Morey & Rouder, 2015; Rouder et al., 2012; Westfall et al., 1997). Accuracy and reaction times (RTs) from the gaze cueing tasks were analysed introducing the factors SOA (2 levels: 200 ms, 700 ms), Face trait (4 levels: HT, LT, HD, LD) and Cue (valid, invalid). Accuracy was overall high in the three experiments (about 95%). Such a high accuracy suggests focusing the analysis on the RTs of correct responses. Nonetheless, the accuracy analyses are reported in Supplementary material Part B for completeness. Only RTs for correct responses were considered, and RTs were filtered to exclude trials with responses faster than 300 ms and slower than 1500 ms. Responses from the rating tasks were analysed introducing the factors Time (2 levels: pre-, post-gaze cueing paradigm), Face trait (4 levels: HT, LT, HD, LD), gaze (2 levels: direct, averted). The model comparison function of Bayesian ANOVA was used to provide information about the relative adequacy of each possible model. Outputs of model comparisons presented in full in Supplementary material report Bayes Factor (BF) as the ratio that contrasts the likelihood of the data fitting under the best model (i.e., the most probable model given the data) with the likelihood of fitting under each alternative model (BF_{01}); thus, higher BF_{01} indicates greater support to the model listed as the best model compared to each of the other listed models. Using the classification scheme from JASP guidelines (van Doorn et al., 2020) for the interpretation of BF, a BF between 3 and 10 suggests moderate evidence and a BF > 10 suggests strong evidence for the best model compared to other listed models. In addition, information on specific effects of interest was obtained by looking at "$BF_{inclusion}$", which is the change from prior to posterior odds for a factor averaged across the models that contain that specific factor (Wagenmakers et al., 2018).

Results

Experiment 1

Gaze cueing effect

Our first question of interest centres on whether perceived trustworthiness or dominance influences the gaze cueing effect at different SOAs. The mean proportion correct overall accuracy in the gaze cueing part of the experiment was 0.95 (SD=0.04), after excluding one participant because of low accuracy (M=0.08); this high overall accuracy suggests to focus on the correct response times and Bayesian analysis provided support to

DOI: 10.1177/17470218211007791

the null model (analyses on accuracy are reported in the online Supplementary material part B; see Table S10). The overall mean of RTs for correct responses was 569 ms (SD=81.37). Bayesian ANOVA showed that the model with the greatest support included the main effects of SOA and Cue, whereas the first model including Face trait as a main effect was 79 times less likely compared to the best model and the first model including the Cue by Face trait interaction was extremely less probable than the best model (BF_{01} = 1741) (see Supplementary Table S1). Looking at the inclusion Bayes factor for each specific effect, there was clear evidence in the data for including SOA ($BF_{inclusion}$ = $3.11e^{+7}$) and Cue ($BF_{inclusion}$ = $7.55e^{+11}$) as predictors in the model, whereas there were evidence suggesting to exclude Face trait from the predictors as a main effect ($BF_{inclusion}$ = 0.005) or in interaction with other factors (SOA x Face traits: $BF_{inclusion}$ = $5.8e^{-4}$, Cue x Face trait: $BF_{inclusion}$ = 0.001, SOA x Cue x Face trait: $BF_{inclusion}$ = $4.21e^{-6}$).

In summary, RTs were faster in trials with 700 ms (M=564, SD=82.07) than 200 ms SOA (M=573, SD=81.49) and faster for valid compared to invalid trials (Figure 2). The Bayesian analysis clearly revealed that RTs were influenced by gaze cue and SOA, but not by facial trustworthiness and dominance.

[Figure 2 about here]

Incidental learning of trustworthiness and dominance

Our second question of interest involves whether incidental learning in the gaze cueing task influences the perceived trustworthiness and dominance of the faces themselves. For this purpose, we compared ratings collected before (pre-) and after (post-) the gaze cueing task. These rating scores are reported in Tables 1 and 2.

A Bayesian repeated measures ANOVA on trustworthiness ratings showed that the best model included the main effects of Face trait and Gaze. The model including only the main effect of Face trait was 1.5 times less probable, whereas the model including the three main effects of Face trait, Gaze and Time was about 7 times less probable (See supplementary Table S2). Inclusion Bayes factors provided overwhelming support for including Face trait as a predictor ($BF_{inclusion}$ = ∞), whereas the support for including Gaze as a predictor was inconclusive ($BF_{inclusion}$ = 0.56) and there was no evidence for including other effects (all $BF_{inclusion}$ < 1). For the Face trait effect the adjusted posterior odds of post-hoc tests showed evidence that all faces differed from

DOI: 10.1177/17470218211007791

each other (posterior odds: HT vs LT = 3.39^{+34}, HT vs HD = $4.86e^{+18}$, HT vs LD = $9.01e^{+16}$, HD vs LD = 359.3, HD vs LT = $8.12e^{+14}$, LD vs LT = $9.77e^{+19}$). As expected, HT and LT faces received the highest and lowest scores in rated trustworthiness, respectively; scores for LD and HD faces were in the middle with higher trustworthiness ratings for the LD than HD faces (see Table 1). Moreover, faces with direct gaze were rated as more trustworthy than faces with averted gaze, although alternative hypothesis for this effect received inconclusive support from the Bayes factor. In Experiment 1 the gaze cues provided by each face were as likely to be valid as invalid, thus ratings were not expected to be influenced by Time (i.e. not expected to differ pre- or post- the gaze cueing task). Indeed, Bayesian analysis revealed that trustworthiness scores were not influenced by Time.

Analysis of dominance ratings showed that the best model included the main effect of Face trait. Models with Face trait and Gaze main effects or Face trait and Time main effects were about 7 times and 11 times less probable than the best model, respectively (see Supplementary Table S3). Inclusion Bayes factors provided overwhelming support for including Face trait as a predictor ($BF_{inclusion} = \infty$), whereas there was no evidence for including other effects (all $BF_{inclusion} < 1$). Post hoc tests showed evidence suggesting equal scores for HD and LT faces (posterior odds = 0.12) and extreme evidence supporting differences in other paired comparisons (posterior odds: HD vs LD = $2.28e^{+38}$, HD vs HT = $3.42e^{+21}$, HT vs LD = 808.34, HT vs LT =$9.7e^{+13}$, LD vs LT = $1.61e^{+30}$). HD and LT faces received the highest scores, followed by HT and LD faces, which received the lowest dominance score as expected (see Table 2). Therefore, the analysis clearly showed that dominance ratings were influenced by Face trait, but not by Time, as was expected when the gaze cues were valid and invalid with equal probability for each face.

[Tables 1 and 2 about here]

Experiment 2

Gaze cueing effect

The mean overall proportion of correct responses in Experiment 2 was 0.96 (SD=0.04); Bayesian analysis provided support to the null model (Table S11 in Supplementary material part B). The overall mean of RTs for correct responses was 549 ms (SD=61.16). Bayesian analysis showed that the model with greatest

DOI: 10.1177/17470218211007791

support included the main effects of SOA and Cue, whereas the first model including Face trait as a main effect was 56 times less likely compared to the best model and the first model including the Cue by Face trait interaction was 3382 times less probable than the best model (see Supplementary Table S4). Inclusion Bayes Factors provided overwhelming support for introducing SOA ($BF_{inclusion}$ = 5.61e^{+6}) and Cue ($BF_{inclusion}$ = 1.07e^{+15}) as predictors in the model, whereas evidence suggested to exclude Face trait as a main effect ($BF_{inclusion}$ = 0.005) or in interaction with other factors (all $BF_{inclusion}$ < 0). Thus, as in Experiment 1, RTs were faster for 700 ms (M=544, SD=65.94) than 200 ms SOA (M=554, SD=57.97) and faster for valid than invalid cues (see Figure 3). The analysis provided clear evidence that RTs were affected by SOA and Cue but no evidence for modulation by trustworthiness and dominance.

[Figure 3 about here]

Incidental learning of trustworthiness and dominance

Analysis of trustworthiness ratings (Table 1) showed that the best model included the main effects of Gaze and Face trait, although there was inconclusive evidence (BF_{01} = 1.24, error % = 2.98) that the best model outperformed a more complex model, which included the three main effects of Gaze, Time and Face trait and the interaction Time x Face trait (see Supplementary Table S5). Inclusion Bayes factors showed that the data strongly supported the inclusion of Face trait ($BF_{inclusion} = \infty$) and Gaze ($BF_{inclusion}$ = 87.45) main factors, whereas the interaction Time x Face trait ($BF_{inclusion}$ = 1.33) and Time main factor ($BF_{inclusion}$ = 0.51) fell in the area of uncertainty. Trustworthiness ratings were higher for direct than averted gaze. Looking at the post hoc tests for the main effect of Face trait, adjusted posterior odds showed that the highest score of HT faces differed from HD (posterior odds = 3.9e^{+11}), LD (posterior odds = 1.41e^{+19}) and LT faces (posterior odds = 1.96e^{+26}). LT faces received the lowest score, which differed from HD (posterior odds = 1.23e^{+11}) and LD faces (posterior odds = 4.68e^{+8}), whereas the data supported the equivalence between LD and HD faces (posterior odds = 0.07), whose scores were in the middle between HT and LT faces. The main analysis provided inconclusive evidence concerning the Time x Face trait interaction. We thus conducted a series of Bayesian paired sample t-tests to explore potential differences between pre- and post-gaze cueing paradigms in trust rating for each face trait. These post hoc analyses have been conducted with an exploratory intent,

DOI: 10.1177/17470218211007791

with the idea of generating potential hypothesis that will deserve future specific experiments. Results showed evidence in support of the null hypothesis for HT (BF_{10} = 0.4, error % = 3.89e^{-6}), LT (BF_{10} = 0.2, error % = 8.64^{-6}) and HD faces (BF_{10} = 0.2, error % = 8.75e^{-6}), whereas the alternative hypothesis was supported for LD faces (BF_{10} = 133.96, error % = 7.74e^{-5}), which decreased in rated trustworthiness in the second assessment (Figure 4).

In Experiment 2, then, the probabilities of valid and invalid cues associated with different faces were manipulated: HT and HD faces provided the majority of valid cues, whereas LT and LD faces provided the majority of invalid cues. Thus, incidental learning from the gaze cueing task was hypothesised, which would lead to differences across pre- and post- ratings. Bayesian analysis clearly confirmed that Face traits influenced trustworthiness ratings. Additionally, despite uncertain evidence at factor level, the incidental learning reduced trustworthiness score for LD faces following the gaze cueing task. This latter result should be taken cautiously. However, our evidence suggests deepening the investigation of LD faces in future studies.

[Figure 4 about here]

Analysis of dominance ratings showed that the best model included the three main effects of Time, Face trait and Gaze, although there was inconclusive evidence (BF_{01} = 1.1, error % = 4.35) that the best model outperformed the model with only two main effects of Face trait and Gaze (see Supplementary Table S6). In fact, the data supported the inclusion of Face trait ($BF_{inclusion}$ = 1.61e^{+15}) and Gaze ($BF_{inclusion}$ = 20.07), but there was no evidence in support, but neither against, for including Time as a predictor ($BF_{inclusion}$ = 0.47). Dominance ratings were higher for faces with central than averted gaze. Post hoc tests showed evidence of no differences between HD and LT faces (posterior odds = 0.04) and extreme evidence supporting differences in other paired comparisons (posterior odds: HD vs LD = 1.04e^{+36}, HD vs HT = 2.94e^{+19}, HT vs LD = 6.05e^{+7}, HT vs LT =2.61e^{+14}, LD vs LT = 8.28e^{+31}). As in Experiment 1, HD and LT faces received the highest scores followed by HT and LD faces, which received the lowest dominance score as expected (see Table 2). Therefore, dominance ratings were modulated by facial traits and gaze direction, but there was no evidence of incidental learning from the gaze cueing task affecting perceived dominance.

DOI: 10.1177/17470218211007791

Experiment 3

Gaze cueing effect

The overall mean proportion of correct responses was 0.96 (SD=0.04); the best model resulting from Bayesian analysis was the null model, although in this case there was inconclusive evidence that the null model outperformed the model with the factor Cue (Table S12 in Supplementary material part B). The overall mean RTs for correct responses was 567 ms (SD = 64.7). Bayesian analysis of RTs for correct responses showed that the model with the greatest support included the main effects of SOA and Cue, whereas the first model including Face trait as a main effect was about 9 times less likely compared to the best model and the first model including Face trait by Cue interaction was about 178 times less likely than the best model (see Supplementary Table S7). Inclusion Bayes Factors provided overwhelming support for introducing SOA ($BF_{inclusion} = 3.22e^{+15}$) and Cue ($BF_{inclusion} = 14717$) as predictors in the model, whereas there was evidence for excluding Face trait as a main effect ($BF_{inclusion} = 0.05$) or in interaction with other factors (SOA x Face traits: $BF_{inclusion} = 0.05$, Cue x Face trait: $BF_{inclusion} = 0.01$, SOA x Cue x Face trait: $BF_{inclusion} = 1e^{-4}$).

In summary, as in the previous experiments, RTs were faster in trials with 700 ms (M=559, SD=64.69) than 200 ms SOA (M=574, SD=65.33) and faster for valid than invalid cues (see Figure 5), whereas there was no evidence supporting the modulation of these effects by face traits.

[Figure 5 about here]

Incidental learning of trustworthiness and dominance

Analysis of trustworthiness ratings revealed that the best model included the main effects of Gaze and Face trait, and the first model including the Time factor was about 11 times less probable than the best model ($BF_{01} = 10.86$, error % = 5.81, see Supplementary Table S8). Inclusion Bayes factors provided weak evidence for including Gaze as a predictor ($BF_{inclusion} = 3.51$) and overwhelming support for including Face trait as a predictor ($BF_{inclusion} = \infty$). Trustworthiness ratings were lower for averted than central gaze (Table 1). Post hoc tests for the main effect of Face trait showed that the highest ratings were for HT faces, which

DOI: 10.1177/17470218211007791

differed from HD (posterior odds = $1.91e^{+6}$), LD (posterior odds = $1.7e^{+14}$), and LT faces (posterior odds = $2.7e^{+25}$); LT faces received the lowest score, which differed from HD (posterior odds = $3.34e^{+14}$) and LD faces (posterior odds = $4.86e^{+7}$), whereas the data supported the equivalence between LD and HD faces (posterior odds = 0.06).

In Experiment 3, probabilities of valid and invalid cues associated with each face were manipulated oppositely to Experiment 2: HT and HD faces provided the majority of invalid cues, whereas LT and LD faces provided the majority of valid cues. Incidental learning from gaze cueing task was hypothesised, but the analysis did not show evidence of differences in trustworthiness ratings pre- and post-gaze cueing. On the other hand, results confirmed the modulation of trustworthiness ratings by gaze and facial traits.

Bayesian repeated measures ANOVA of dominance ratings showed that the best model included the main effects of Time, Face trait and Gaze factors, although there was relatively weak evidence that it outperformed the model with only Face trait and Gaze factors (BF_{01} = 3.34, error % = 4.137) (see Supplementary Table S9). Looking at the inclusion Bayes factor there was clear evidence in the data for including Face trait as a predictor in the model ($BF_{inclusion} = \infty$), weak evidence for including Gaze ($BF_{inclusion}$ = 3.43), and inconclusive evidence toward any direction for including Time ($BF_{inclusion}$ = 1.48). Dominance ratings were higher for faces with direct than averted gaze and tended to increase following the gaze cueing paradigm compared to dominance ratings collected prior to the gaze cueing task (see Table 2). Post hoc tests for the main effect of Face trait showed weak evidence for the equivalence between HD and LT faces (posterior odds = 0.28) and extreme evidence supporting differences in other paired comparisons (posterior odds: HD vs LD = $1.3e^{+42}$, HD vs HT = $4.73e^{+21}$, HT vs LD = $2.04e^{+7}$, HT vs LT =$9.15e^{+14}$, LD vs LT = $4.65e^{+35}$). As in the previous experiments, HD and LT faces received the highest scores, followed by HT and LD faces, which received the lowest dominance score as expected (see Table 2). Therefore, the analysis showed that dominance ratings were influenced by gaze and face traits and suggested, albeit with weak evidence from the data, a modulation after performing the gaze cueing task.

Discussion

This study investigated the modulatory impact of face trustworthiness and dominance on the gaze cueing effect, and in addition, assessed incidental learning of social judgments following the gaze cueing paradigm

DOI: 10.1177/17470218211007791

when different contingencies of valid and invalid cues were associated with different faces. For the gaze cueing effect itself, findings were clear. Across three experiments, the results consistently showed that the gaze cueing effect was not modulated by perceived social traits of the faces. Indeed, this was supported by strong evidence from Bayesian analyses for not including Face trait, either as a main effect or in interaction with other factors, in the data modelling of the three gaze cueing experiments. In particular, results found no evidence of interactions between Face trait and Cue validity; thus the orienting of attention in response to gaze cues was not modulated by different social traits. On the other hand, results on incidental learning were less clear, as data provided weak or inconclusive evidence for the influence of gaze cueing on subsequent social judgments, with different results depending on the social traits of faces associated with different proportions of valid or invalid cues.

First, we will discuss the data from the gaze cueing paradigm. In Experiment 1, faces were associated with valid and invalid cues with equal probability. In Experiment 2 HT and HD faces provided 70% of valid cues, whereas LT and LD faces provided 70% of invalid cues; thus contingencies were manipulated in line with the hypothesis that highly trustworthy and dominant faces would increase gaze-mediated orienting of attention. In contrast with the latter expectation, in Experiment 3 highly trustworthy and dominant faces had greater probability of deceiving, since HT and HD faces were associated with 70% of invalid cues and LT and LD were associated with 70% of valid cues. Crucially, the gaze cueing effect was consistent across the three experiments, with faster RTs for valid than invalid trials regardless of the face trait manipulation.

We also manipulated SOA in the gaze cueing paradigm, with the cue presented 200 ms or 700 ms before target onset, to ascertain whether a longer SOA might create a stronger influence of the face's social characteristics. Although RTs were consistently faster for the longer SOA, there were no interactions between the Cue, SOA and Face trait factors; instead, Bayesian analyses showed that the best predictive models for all three experiments included only main effects of Cue and SOA. These results replicate previous findings showing that the gaze cueing effect is a robust phenomenon, with reduced response latencies as SOA increased likely to reflect a general warning effect (Friesen and Kingstone, 1998; Driver et al., 1999; Frischen et al., 2007).

The lack of influence of trustworthiness and dominance on the basic gaze cueing effect is of particular interest because a number of other studies have shown that consistent impressions of traits such as

DOI: 10.1177/17470218211007791

trustworthiness or dominance can be formed to presentations as short as 100 ms (Ballew & Todorov, 2007; Bar et al., 2006; Olivola & Todorov, 2010; South Palomares & Young, 2018; Todorov et al., 2009; Willis & Todorov, 2006), which is more rapid than the short SOA used here. However, it needs to be noted that the fact that consistent impressions can be created to 100 ms presentation does not imply that such impressions are themselves necessarily formed within 100 ms and affect attentional behaviour. In contrast to our findings, influences of social modulators on the gaze cueing effect have been reported in some previous studies which varied face age, dominance or trustworthiness, or manipulated the social status of stimuli with biographical descriptions of characters (Carraro et al., 2017; Ciardo et al., 2014; Jones et al., 2010; Ohlsen et al., 2013; Petrican et al., 2013; Süßenbach & Schönbrodt, 2014). However, social influence was not significant in other studies (King et al., 2011; Strachan et al., 2017). Although the sample size of our experiments was determined on the basis of previous literature (Ciardo et al., 2014; Jones et al., 2010; Ohlsen et al., 2013; King et al., 2011; Strachan et al., 2017), the use of Bayesian statistics also helps by assessing the strength of the evidence. Thus, we additionally adopted a sequential analysis to assesses the possibility of having run an underpowered study. The sequential analysis (see supplementary material) showed a general agreement excluding the role of Face trait even with very small sample sizes (less than 10 participants). Differently, the uncertainty about the SOA x Cue interaction has to be acknowledged. Future studies, more powered, may shed light on this additional element.

To the best of our knowledge, ours is the first study assessing the impact of trustworthiness and dominance dimensions across the same gaze cueing paradigm using stereotypical faces created with a photograph averaging technique (Sutherland et al., 2013; Sutherland et al., 2017). Therefore, stimuli were not constrained in cues from the manipulated dimensions; instead they varied in naturally occurring features related to each social trait, such as age or emotional expression. This increases the ecological validity of results, although we cannot then isolate the impact of each of these face properties. The absence of gaze cueing modulations related to social traits of faces was consistent across our three experiments and strongly supported by the Bayesian statistical approach. Thus, this study confirms that attentional shifts in the gaze cueing paradigm reflected a relatively automatic process, which was unaffected by social traits of faces. Recent reviews (Dalmaso et al., 2020a; Barbato et al., 2020) highlighted controversial results concerning social modulators of the gaze cueing effect, in particular when complex and multifactorial social constructs

DOI: 10.1177/17470218211007791

were explored, and suggested that the impact of social variables could depend on the relevance of such variables for the task or the observer (see also Ricciardelli et al., 2013; Ciardo et al., 2014). We can interpret our results as in line with this idea, since the experimental task of discriminating target letters did not induce the participants to consider trustworthiness and dominance of face stimuli as relevant features for the task itself. It is also worth noting that previous studies reported modulations of the gaze cueing effect related to the features of the paradigms, which presented different stimulus categories intermixed across trials or separated in blocked conditions (Dalmaso et al., 2020b; Pavan et al., 2011; Zhao et al., 2014). In our experiments, stimuli varying in two social dimensions (i.e. high/low dominance and trustworthiness) were randomly presented across trials, and this could have reduced the impact of social traits on attention orienting. This speculative interpretation could be tested in future research, which might address the hypothesis that a more consistent presentation of stimuli belonging to the same category could affect the impact of such category on gaze cueing, possibly increasing its salience.

Next we turn to consider the ratings of trustworthiness and dominance. Here, the data confirmed that the stimuli varied as expected in their intended dimensions across the three experiments: HT and LT faces received the highest and the lowest trustworthiness ratings, respectively, whereas HD and LD faces received the highest and the lowest dominance ratings, respectively. These differences were already apparent when ratings were collected before participants carried out the gaze cueing task, confirming the validity of the stimuli themselves. As already noted, the value of the computer-averaged stimuli is that they capture as many as possible of the naturally occurring cues to a particular trait. By taking widely differing everyday 'ambient' images (cf. Jenkins et al., 2011) and averaging those seen as high on a particular trait, cues consistent with that trait remain in the averaged image and irrelevant variation is averaged away (Sutherland et al., 2017). This constitutes a data-driven approach that carries no assumptions about which cues (or cue combinations) signal a particular trait and that can capture natural covariation between the cues themselves (Sutherland et al., 2017; Young, 2018). However, it is worth noting that in consequence, the dimensions of trustworthiness and dominance were not orthogonal. Instead, variations of high and low dominance faces along the trustworthiness dimension differed from variations of high and low trustworthy faces along the dominance dimension. Indeed, HD and LD faces did not vary consistently in perceived trustworthiness: they received trustworthiness ratings in the middle of the range between HT and LT faces across the three

DOI: 10.1177/17470218211007791

experiments, with Bayesian analyses supporting higher scores for LD than HD faces only in Experiment 1, but supporting the absence of differences in Experiments 2 and 3. On the other hand, HT and LT faces consistently differed in dominance scores, with HT faces being perceived as less dominant than LT faces, with scores similar to HD faces for the latter. This evidence of some intercorrelation between trustworthiness and dominance dimensions reflects the technique used for creating the factor model underlying the stimuli, which was data-driven and did not force the dimensions to be orthogonal; instead it retained naturally occurring covariation between such different features of faces as emotional expressions, age, and skin tone (Sutherland et al., 2013). We note that whilst some computer-generated images have been designed to model trustworthiness and dominance as orthogonal (Oosterhof & Todorov, 2008), other data-driven techniques using photographs do also show some overlap between trustworthiness and dominance dimensions (Oosterhof & Todorov, 2008).

In line with previous studies (Main et al., 2009; Kaisler & Leder, 2016) the main effect of gaze direction on perceived social characteristics of the faces was present in our experiments. Although with different rate of support from Bayesian statistics (i.e. inconclusive in Experiment 1, strong in Experiment 2, and weak in Experiment 3), data from the three experiments showed higher dominance and trustworthiness scores for faces with gaze directed to participants than when gaze was averted leftwards or rightwards. Changes between pre- and post-gaze cueing were hypothesised for Experiments 2 and 3, in which different faces provided different proportions of valid and invalid cues. In particular, we expected a Time by Face trait interaction depending on the contingency associated with each type of face, which would indicate incidental learning from gaze behaviour (Strachan et al., 2017). This hypothesis received only partial support, within the limit of uncertainty in Bayesian analysis, from data of Experiment 2, which showed that trustworthiness ratings of LD faces were somewhat reduced following the gaze cueing paradigm; LD faces provided only 30% of valid cues in Experiment 2, as well as LT faces. However, while LT faces received the lowest trustworthiness score already in the first rating session, in line with the stereotypical manipulation of the stimuli, LD faces were evaluated with intermediate scores on the trustworthiness scale (M=4.58) at the beginning of the experiment, but the participants reduced trustworthiness judgments in the second rating session (M=4.0) following the gaze cueing paradigm in which LD faces deceived on most trials. A different effect emerged in Experiment 3, in which contingencies were oppositely manipulated, namely HD and HT

DOI: 10.1177/17470218211007791

faces deceived, whereas LD and LT faces cooperated. In this case, there was no influence of the gaze cueing paradigm on trustworthiness judgments, whereas dominance scores increased in the second rating session. Again, Bayesian statistics suggested a cautious discussion on this result, as the best model, which included Time main effect, received weak support compared to a simpler model with only Face trait and Gaze modulations. Moreover, the analysis showed a main effect of Time without interaction with Face trait. Thus, the increased dominance score was not related to the different gaze behaviour of the face stimuli. These results might suggest that incidental learning from gaze behaviour had a different impact on the evaluation of trustworthiness and dominance dimensions depending on the association between stereotypical traits and face behaviour in the gaze cueing paradigm. Indeed, trustworthiness judgments decreased only for LD faces who deceived in Experiment 2. On the other hand, when expectations based on stereotypical trait associations were violated, as in Experiment 3, incidental learning was not so strong as to affect trustworthiness judgments, but all faces were perceived as more dominant following the gaze cueing paradigm. The latter result could be interpreted in the context of studies showing that facial features related to dominance can predict deception (Stirrat & Perrett, 2010; Haselhuhn & Wong, 2012). Indeed, we could speculate that gaze behaviour that contradicts expectations is interpreted by the participants as a non-cooperative or unethical behaviour, and this increases the perceived dominance of faces.

Previous studies have reported incidental learning of trustworthiness following the gaze cueing paradigm with fixed contingencies of valid and invalid cues associated to different faces and showed that the effect of gaze behaviour on trust learning was modulated by facial ethnicity or emotional expressions (Bayliss et al., 2009; Bayliss & Tipper 2006; Manssuer et al., 2015; 2016; Strachan et al., 2017). However, the effect of the gaze cueing paradigm on dominance evaluation has not previously been assessed. Our experiments thus add preliminary evidence to this literature, suggesting that incidental learning of trustworthiness and dominance following gaze cueing might be modulated in different ways depending on contingencies in gaze behaviour and the baseline perception of these social dimensions in face features.

In conclusion, our study adds new evidence on the complex and reciprocal influences between attention orientation from eye gaze and social face traits. The Bayesian statistical approach provided consistent support to an absence of modulation of the gaze cueing effect from stereotypical faces with high or low levels of perceived trustworthiness and dominance. From this we conclude that gaze cueing is relatively

DOI: 10.1177/17470218211007791

automatically generated, without reference to many social properties of the faces when (as in this case) they are task-irrelevant. On the other hand, we obtained uncertain results concerning the influence of gaze cueing behaviour on subsequent social judgments through incidental learning, with different modulations related to face traits. Since Bayesian statistics highlighted uncertainty on the latter, further studies are needed to clarify the direction of the data in favour or against the hypothesis of incidental learning from gaze cueing. These results confirm the value of using ecologically created stimuli to investigate social face features affecting the gaze cueing paradigm and highlight that exploring the relationship between context, perceivers and stimuli is crucial to understanding the mechanisms involved in the human orientation of attention in social environments (Ricciardelli et al., 2013; Ciardo et al., 2014; Dalmaso et al., 2020a).

Supplementary Material

The Supplementary Material is available at: qjep.sagepub.com

Funding Acknowledgments

This research was funded by a grant (Fondo di Ateneo) from the University of Milano-Bicocca to PR.

Declaration of conflicting interests

The Authors declare that there is no conflict of interest.

References

Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences, 104(46)*, 17948–53. http://doi.org/10.1073/pnas.0705435104.

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

Barbato, M., Almulla, A. A., & Marotta, A. (2020). The Effect of Trust on Gaze-Mediated Attentional Orienting. *Frontiers in Psychology*, 11, 1554. https://doi.org/10.3389/fpsyg.2020.01554

DOI: 10.1177/17470218211007791

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. doi: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. doi:10.1111/j.1467-9280.2006.01737.x

Bonifacci, P., Ricciardelli, P., Lugli, L., & Pellicano, A. (2008). Emotional Attention: Effects of Emotion and Gaze Direction on Overt Orienting of Visual Attention. *Cognitive Processing*, 9, 127-135. doi: 10.1007/s10339-007-0198-3.

Evans, A. M., & Revelle, W. (2008). Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality*, 42(6), 1585–1593. https://doi.org/10.1016/j.jrp.2008.07.011

Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I., & Young, A. W. (2002). Reading the mind from eye gaze. *Neuropsychologia*, 40(8), 1129–1138. https://doi.org/10.1016/S0028-3932(02)00008-8

Carraro, L., Dalmaso, M., Castelli, L., Galfano, G., Bobbio, A., & Mantovani, G. (2017). The appeal of the devil's eye: social evaluation affects social attention. *Cognitive Processing*, 18(1), 97–103. https://doi.org/10.1007/s10339-016-0785-2

Ciardo, F., Angelis, J. De, Marino, B. F. M., Actis-grosso, R., & Ricciardelli, P. (2021). Acta Psychologica Social categorisation and joint attention : Interacting effects of age , sex , and social status. *Acta Psychologica*, 212, 103223. https://doi.org/10.1016/j.actpsy.2020.103223

Ciardo, F., Marino, B. F., Actis-Grosso, R., Rossetti, A., & Ricciardelli, P. (2014). Face age modulates gaze following in young adults. *Scientific Reports*, 4(4746). doi:10.1038/srep04746

Dalmaso, M., Castelli, L., & Galfano, G. (2020a). Social modulators of gaze-mediated orienting of attention: A review. *Psychonomic Bulletin & Review*, 27, 833-855. https://doi.org/10.3758/s13423-020-01730-x

Dalmaso, M., Alessi, G., Castelli, L., & Galfano, G. (2020b). Eye contact boosts the reflexive component of overt gaze following. *Scientific Reports,* 10, 4777. https://doi.org/10.1038/s41598-020-61619-6

Dalmaso, M., Pavan, G., Castelli, L., & Galfano, G. (2012). Social status gates social attention in humans. *Biology Letters*, 8(3), 450–452. https://doi.org/10.1098/rsbl.2011.0881

DOI: 10.1177/17470218211007791

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

Ewing, L., Rhodes, G., Pellicano, E. (2010). Have you got the look? Gaze direction affects judgements of facial attractiveness. *Visual Cognition*, 18(3), 321-330.

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

Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional Effects of Counterpredictive Gaze and Arrow Cues. *Journal of Experimental Psychology: Human Perception and Performance,* 30(2), 319–329. https://doi.org/10.1037/0096-1523.30.2.319

Frischen, A., Bayliss, A. P., & Steven P., T. (2007). Gaze Cueing of Attention: Visual Attention, Social Cognition, and Individual Differences Alexandra. *Psychological Bulletin*, 133(4), 694–724. https://doi.org/10.1037/0033-2909.133.4.694.

Frischen, A., & Tipper, S. P. (2006). Long-term gaze cueing effects: Evidence for retrieval of prior states of attention from memory. Visual Cognition, 14(3), 351–364. https://doi.org/10.1080/13506280544000192

Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: Facial structure predicts unethical behaviour. *Proceedings of the Royal Society B: Biological Sciences,* 279(1728), 571–576. https://doi.org/10.1098/rspb.2011.1193

Hietanen, J. K., & Leppänen, J. M. (2003). Does Facial Expression Affect Attention Orienting by Gaze Direction Cues? *Journal of Experimental Psychology: Human Perception and Performance*, 29(6), 1228–1243. https://doi.org/10.1037/0096-1523.29.6.1228

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition, 121(3)*, 313-323. https://doi.org/10.1016/j.cognition.2011.08.001

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

DOI: 10.1177/17470218211007791

Jones, B. C., Main, J. C., Little, A. C., & DeBruine, L. M. (2011). Further evidence that facial cues of dominance modulate gaze cuing in human observers. *Swiss Journal of Psychology*, 70(4), 193–197. https://doi.org/10.1024/1421-0185/a000056

Kaisler, R. E., & Leder, H. (2016). Trusting the looks of others: gaze effects of faces in social settings. *Perception,* 45 (8), 875-892.

King, D., Rowe, A., & Leonards, U. (2011). I trust you; hence I like the things you look at: Gaze cueing and sender trustworthiness influence object evaluation. *Social Cognition*, 29,476–485.

Ly, A., Verhagen, A. J. & Wagenmakers, E.-J. (2016). Harold Jeffreys's Default Bayes Factor Hypothesis Tests: Explanation, Extension, and Application in Psychology. *Journal of Mathematical Psychology*, 72, 19-32.

Ly, A., Marsman, M., Wagenmakers, E.-J. (2018). Analytic Posteriors for Pearson's Correlation Coefficient. *Statistica Neerlandica*, 72(1), 4-13.

Main, J. C., Jones, B. C., DeBruine, L. M., Little, A. C. (2009). Integrating gaze direction and sexual dimorphism of face shape when perceiving the dominance of others. *Perception,* 38(9), 1275-1283.

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

Mason, M. F., Tatkow, E. P., Macrae, C. N. (2005). The look of love. Gaze shifts and person perception. *Psychological Science,* 16 (3), 236-239.

Mattavelli, G., Andrews, T. J., Asghar, A. U., Towler, J., & Young, A. W. (2012). Response of face-selective brain regions to trustworthiness and gender of faces. *Neuropsychologia, 50(9)*, 2205-2211. doi: 10.1016/j.neuropsychologia.2012.05.024

Milders, M., Hietanen, J. K., Leppänen, J. M., Braun, M. (2011). Detection of emotional faces is modulated by the direction of eye gaze*.* *Emotion*, 11 (6), 1456-1461.

Morey, R. D., & Rouder, J. N. (2015). BayesFactor 0.9.11-1. Comprehensive R Archive Network.

DOI: 10.1177/17470218211007791

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science,* 16(5), 269–274. https://doi.org/10.1111/j.1467-8721.2007.00518.x

Oh, D., Dotsch, R., Porter, J., & Todorov, A. (2020). Gender biases in impressions from faces: Empirical studies and computational models. *Journal of Experimental Psychology: General*, 149(2), 323–342. https://doi.org/10.1037/xge0000638

Ohlsen, G., van Zoest, W., & van Vugt, M. (2013). Gender and facial dominance in gaze cuing: Emotional context matters in the eyes that we follow. *PLoS ONE*, 8, e59471.

Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior, 34(2),* 83–110. http://doi.org/10.1007/s10919-009-0082-1

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America,* 105(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

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

Pecchinenda, A., Pes, M., Ferlazzo, F., & Zoccolotti, P. (2008). The combined effect of gaze direction and facial expression on cueing spatial attention. *Emotion,* 8,628–634.

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. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences,* 68(2), 215–223. doi:10.1093/geronb/ gbs064

Ricciardelli, P., Carcagno, S., Vallar, G., & Bricolo, E. (2013). Is gaze following purely reflexive or goal-directed instead? Revisiting the automaticity of orienting attention by gaze cues. *Experimental Brain Research*, 224(1), 93-106. doi: 10.1007/s00221-012-3291-5.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56, 356-374.

Sander, D., Grandjean, D., Kaiser, S., Wehrle, T., Scherer, K. R. (2007). Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *European Journal of Cognitive Psychology,* 19, 470-480.

DOI: 10.1177/17470218211007791

South Palomares, J. K., & Young, A. W. (2018). Facial first impressions of partner preference traits: trustworthiness, status, and attractiveness. *Social Psychological and Personality Science, 9(8)*, 990-1000. https://doi.org/10.1177/1948550617732388

Sprengelmeyer, R., Young, A. W, Baldas, E-M., Ratheiser, I., Sutherland, C. A. M., Müller, H-P., Grön, G., Süssmuth, S. D., Landwehrmeyer, G. B., & Orth, M. (2016). The neuropsychology of first impressions: evidence from Huntington's disease. *Cortex, 85*, 100-115. doi: 10.1016/j.cortex.2016.10.006

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: male facial width and trustworthiness. *Psychological Science*, 21(3), 349–354. https://doi.org/10.1177/0956797610362647

Strachan, J.W., Kirkham, A. J., Manssuer, L. R., Over, H., &Tipper, S. P. (2017). Incidental learning of trust from eye-gaze: Effects of race and facial trustworthiness. *Visual Cognition*, 25,802–814

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. doi:10.1080/20445911.2014.945457

Sutherland, C. A. M., Rhodes, G., & Young, A. W. (2017). Facial image manipulation: a tool for investigating social perception. *Social Psychological and Personality Science, 8(5)*, 538-551. https://doi.org/10.1177/1948550617697176

Sutherland, C. A. M., Young, A. W., Mootz, C. A., & Oldmeadow, J. A. (2015). Face gender and stereotypicality influence facial trait evaluation: Counter-stereotypical female faces are negatively evaluated. *British Journal of Psychology,* 106(2), 186–208. https://doi.org/10.1111/bjop.12085

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

Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton, NJ: Princeton University Press.

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

Todorov, A., & Porter, J. M. (2014). Misleading first impressions: Different for different facial images of the same person. *Psychological Science, 25(7)*, 1404-1417. https://doi.org/10.1177/0956797614532474

DOI: 10.1177/17470218211007791

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences,* 12(12), 455–460. https://doi.org/10.1016/j.tics.2008.10.001

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Science*s, 28(5), 675–691. https://doi.org/10.1017/S0140525X05000129

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., … Wagenmakers, E. J. (2020). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin and Review*. https://doi.org/10.3758/s13423-020-01798-5

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, E3353--E3361. https://doi.org/10.1073/pnas.1409860111

Wagenmakers, E.-J., Maarten, E. W., Tahira, M., Alexander, J., Verhagen, J., Love, J., … Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review,* 25, 35–57.

Westfall, P. H., Johnson, W. O., & Utts, J. M. (1997). A Bayesian perspective on the Bonferroni adjustment. *Biometrika*, 84, 419-427.

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. http://doi.org/10.1111/j.1467-9280.2006.01750.x

Young, A. W. (2018). Faces, people and the brain: The 45th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology, 71(3)*, 569-594. https://doi.org/10.1177/1747021817740275

Zhao, S., Uono, S., Yoshimura, S., & Toichi, M. (2014). Attention orienting by eye gaze and arrows reveals flexibility to environmental changes. *Acta Psychologica*, 150, 100–105. https://doi.org/10.1016/j.actpsy.2014.05.003

DOI: 10.1177/17470218211007791

Figure captions

Figure 1. Timeline of one trial in the Gaze cueing paradigm. In the example a LD male face provides a valid cue to the target letter L. Unmodified (direct gaze) stimuli from Sutherland et al. (2015) are depicted at the right-top of the figure; gaze cue stimuli are modified versions of these with gaze shifted leftward or rightward. HT=high trust; LT=low trust; HD=high dominance; LD=low dominance.

Figure 2. Gaze cueing effect in Experiment 1. RTs for invalid and valid cues **for each face trait and SOA** are depicted, bars represent standard error of the means.

Figure 3. Gaze cueing effect in Experiment 2. RTs for invalid and valid cues **for each face trait and SOA** are depicted, bars represent standard error of the means.

Figure 4. Trustworthiness rating scores in Experiment 2 for ratings collected before (pre) and after (post) participants completed the gaze cueing task. Bars represent means standard errors. * indicates BF_{10} supporting alternative hypothesis.

Figure 5. Gaze cueing effect in Experiment 3. RTs for invalid and valid cues **for each face trait and SOA** are depicted, bars represent standard error of the means.

DOI: 10.1177/17470218211007791

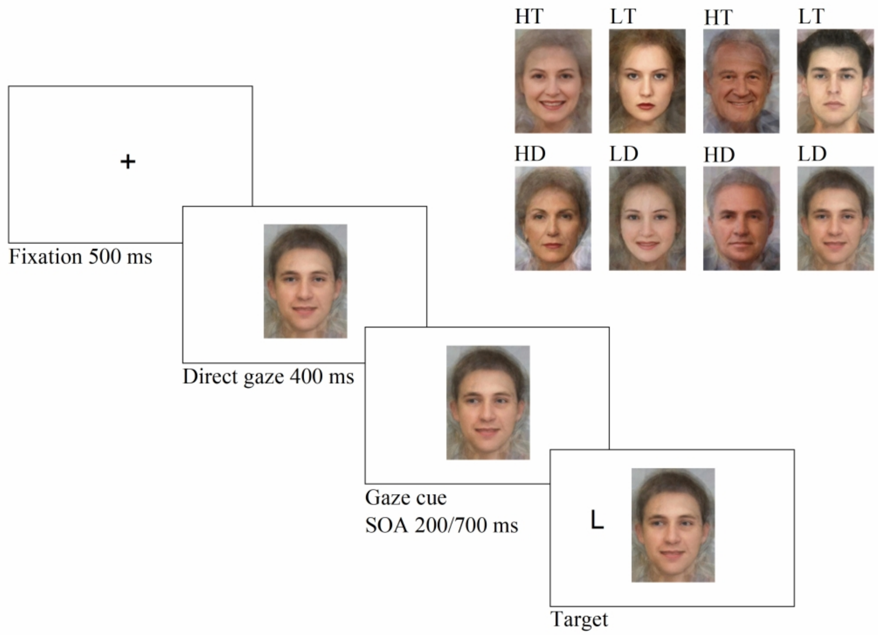

Figure 1. Timeline of one trial in the Gaze cueing paradigm. In the example a LD male face provides a valid cue to the target letter L. Unmodified (direct gaze) stimuli from Sutherland et al. (2015, Figure 3) are depicted at the right-top of the figure; gaze cue stimuli are modified versions of these with gaze shifted leftward or rightward. HT=high trust; LT=low trust; HD=high dominance; LD=low dominance.

287x201mm (300 x 300 DPI)

DOI: 10.1177/17470218211007791

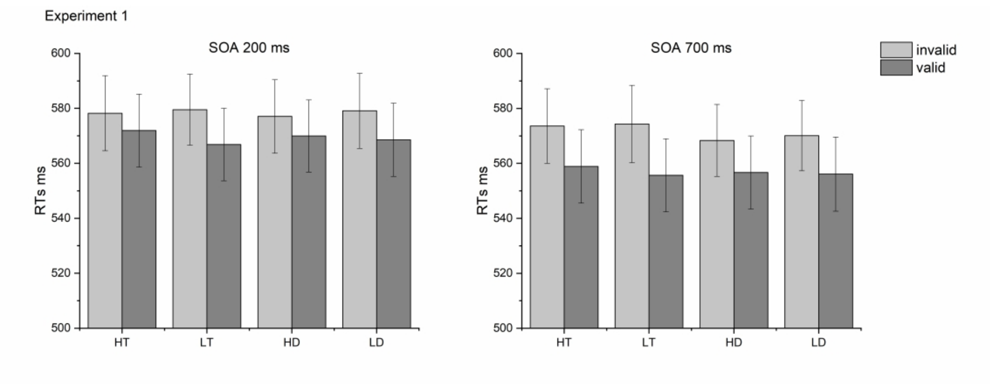

Figure 2. Gaze cueing effect in Experiment 1. RTs for invalid and valid cues for each face trait and SOA are depicted, bars represent standard error of the means.

544x208mm (300 x 300 DPI)

DOI: 10.1177/17470218211007791

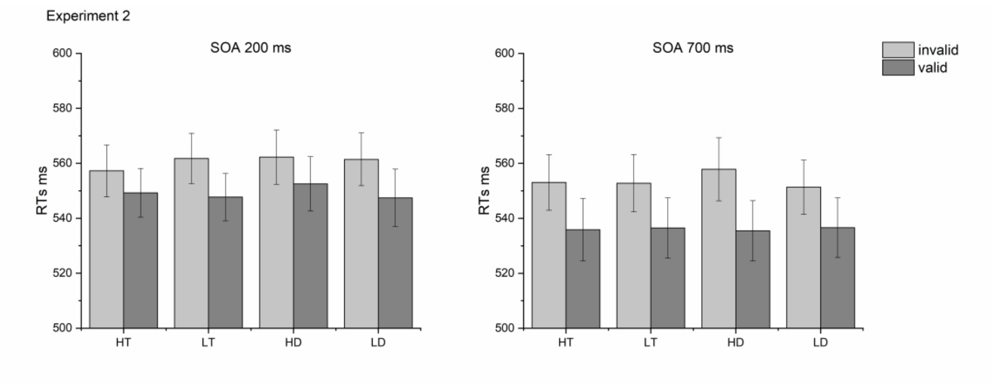

Figure 3. Gaze cueing effect in Experiment 2. RTs for invalid and valid cues for each face trait and SOA are depicted, bars represent standard error of the means.

544x208mm (300 x 300 DPI)

DOI: 10.1177/17470218211007791

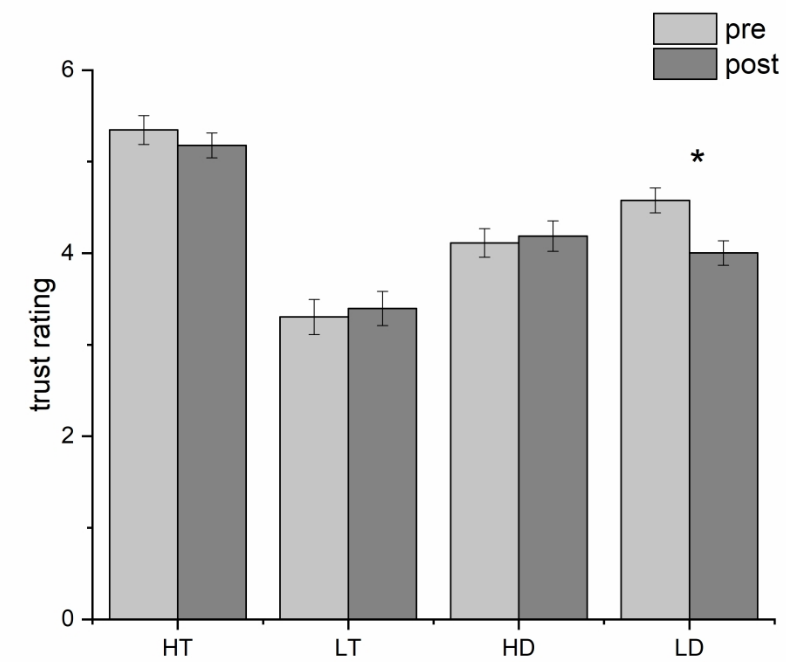

Figure 4. Trustworthiness rating scores in Experiment 2 for ratings collected before (pre) and after (post) participants completed the gaze cueing task. Bars represent means standard errors. * indicates BF10 supporting alternative hypothesis.

272x208mm (300 x 300 DPI)

DOI: 10.1177/17470218211007791

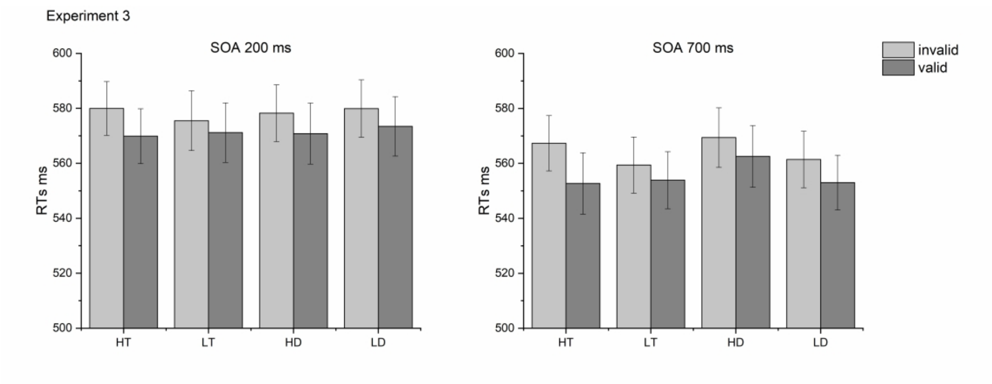

Figure 5. Gaze cueing effect in Experiment 3. RTs for invalid and valid cues for each face trait and SOA are depicted, bars represent standard error of the means.

544x208mm (300 x 300 DPI)

DOI: 10.1177/17470218211007791

Table 1. Means and standard deviations (in brackets) of trustworthiness ratings for faces with central or averted gaze collected before (pre) and after (post) participants completed the gaze cueing task in the three experiments.

		HT		LT		HD		LD	
		central	**averted**	**central**	**averted**	**central**	**averted**	**central**	**averted**
Exp 1	**pre**	5.51 (1.32)	5.21 (1.13)	2.73 (0.94)	2.78 (0.96)	4.22 (1.27)	3.80 (1.07)	4.67 (1.12)	4.32 (1.06)
	post	5.21 (1.28)	5.30 (1.10)	3.03 (1.26)	2.77 (1.19)	3.85 (1.25)	3.76 (1.24)	4.51 (1.26)	4.16 (0.98)
Exp 2	**pre**	5.50 (1.08)	5.19 (1.11)	3.46 (1.31)	3.14 (1.21)	4.40 (1.15)	3.83 (1.08)	4.71 (1.06)	4.44 (0.91)
	post	5.28 (0.91)	5.08 (0.98)	3.46 (1.36)	3.33 (1.09)	4.36 (1.23)	4.01 (1.06)	4.23 (1.08)	3.78 (0.98)
Exp 3	**pre**	5.32 (1.04)	5.02 (0.95)	3.28 (1.11)	3.30 (1.08)	4.66 (1.16)	4.07 (1.15)	4.50 (1.12)	4.13 (1.11)
	post	5.06 (1.22)	5.05 (1.03)	3.54 (1.09)	3.51 (1.07)	4.57 (1.27)	4.28 (1.21)	4.40 (1.03)	4.03 (0.99)

DOI: 10.1177/17470218211007791

Table 2. Means and standard deviations (in brackets) of dominance ratings for faces with central or averted gaze collected before (pre) and after (post) participants completed the gaze cueing task in the three experiments.

		HT		LT		HD		LD	
		central	**averted**	**central**	**averted**	**central**	**averted**	**central**	**averted**
Exp 1	**pre**	3.63 (1.25)	3.55 (1.24)	5.00 (1.35)	4.72 (1.28)	5.23 (1.41)	5.01 (1.00)	3.08 (1.06)	3.00 (0.91)
	post	3.15 (1.14)	3.35 (1.06)	5.21 (1.33)	5.13 (1.29)	5.19 (1.14)	5.21 (1.01)	3.03 (0.78)	2.93 (0.73)
Exp 2	**pre**	3.74 (1.17)	3.48 (1.04)	5.21 (1.06)	4.86 (1.21)	5.20 (0.98)	4.73 (1.09)	3.06 (0.95)	2.76 (0.89)
	post	3.69 (1.36)	3.65 (1.13)	5.29 (1.33)	5.14 (1.27)	5.44 (1.09)	5.05 (1.01)	3.29 (1.11)	2.93 (0.90)
Exp 3	**pre**	3.72 (1.27)	3.52 (1.19)	5.32 (0.84)	4.71 (1.04)	5.11 (1.06)	4.92 (0.96)	3.15 (0.92)	3.01 (0.83)
	post	3.88 (1.45)	3.73 (1.20)	5.17 (1.22)	5.10 (1.15)	5.67 (0.92)	5.34 (0.90)	3.24 (1.02)	3.03 (0.90)

DOI: 10.1177/17470218211007791