



Not to follow because of distrust: perceived trust modulates the gaze cueing effect

Xiaowei Ding¹ · Jing Gan¹ · Luzi Xu¹ · Xiaozhi Zhou¹ · Ding-guo Gao¹ · Yanliang Sun²

Received: 5 June 2023 / Accepted: 27 June 2024

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

In social life, people spontaneously form stable trustworthiness impressions from faces. However, the precise role of extracting trustworthiness information remains unclear. This study aims to elucidate whether discerning facial trustworthiness influences social interactions. Specifically, it explores the gaze cueing effect (GCE), wherein individuals exhibit quicker responses to targets appearing in the direction of gaze compared to other locations. Given conflicting perspectives in existing literature regarding the potential modulation of trustworthiness on the GCE, two plausible hypotheses are proposed to explain divergent result patterns. *The reflexive hypothesis* posits that the GCE operates automatically. In contrast, *the flexible hypothesis* underscores the potential modulatory role of trustworthiness in the GCE. To provide a comprehensive understanding of whether trustworthiness modulates the GCE, we employed face stimuli incorporating trustworthiness information within Posner's cue-target task. The findings of Experiment 1 revealed that the perception of trustworthiness indeed influenced the GCE. Specifically, when facial stimuli were perceived as trustworthy, they elicited a more pronounced GCE compared to untrustworthy stimuli. This modulation effect was replicated using a different stimulus set in Experiment 2. In Experiment 3, we employed the same stimuli as in Experiment 2, setting the trustworthiness information to baseline as a control experiment. The results demonstrated that the trustworthiness modulation effect disappeared, indicating its specificity to the trustworthiness attribute of the stimuli rather than other characteristics. Collectively, these findings lend support to the flexible hypothesis, highlighting that the extraction of trustworthiness information plays a pivotal role in modulating the GCE, consequently influencing social interactions.

Xiaowei Ding and Jing Gan contributed equally to this work and share first authorship.

✉ Ding-guo Gao
edsgao@mail.sysu.edu.cn

✉ Yanliang Sun
yanliangsun@126.com

Xiaowei Ding
dingxw3@mail.sysu.edu.cn

Jing Gan
ganj@mail2.sysu.edu.cn

Luzi Xu
l.xu2@uu.nl

Xiaozhi Zhou
zcx020@126.com

¹ Department of Psychology, Guangdong Provincial Key Laboratory of Social Cognitive Neuroscience and Mental Health, Sun Yat-sen University, Guangzhou, People's Republic of China

² School of Psychology, Shandong Normal University, Jinan 250358, People's Republic of China

Introduction

Following the gaze direction of others plays a crucial role in achieving a fundamental aspect of social interaction, which involves gaining insight into the motivations and needs of individuals (Argyle & Cook, 1976; Clifford & Palmer, 2018). This is primarily due to the wealth of information typically associated with the gaze direction exhibited by people during social interactions (see Emery, 2000; Kleinke, 1986). The phenomenon referred to as the gaze cueing effect (GCE) illustrates the propensity of individuals to shift their attention to the same spatial location as a person they observe looking in a particular direction (Driver et al., 1999; Friesen & Kingstone, 1998; Langton & Bruce, 1999). Numerous studies employing variations of a cue-target paradigm have consistently reported the presence of a GCE (see, e.g., Driver et al., 1999; Friesen & Kingstone, 1998; Kingstone et al., 2019). In fact, the concept of joint attention prompted by GCE holds significant importance in social cognition research (for a review, see Barbato et al., 2020;

Bayliss et al., 2006; Capozzi et al., 2018; Dalmaso et al., 2020; Tomasello, 1995; McKay et al., 2021; Mundy & Newell, 2007). In a seminar work by Bayliss and Tipper (2006), gaze predictability was manipulated by having unpredictable faces alternate between valid and invalid trials, while predictive faces consistently either looked at (valid) or looked away (invalid) from the target. The study found that gaze-evoked attention shifts can influence an observer's evaluation of that person's character (e.g., trustworthiness). This underscores the significance of gaze behavior in social groups and interactions between individuals. Previous studies, exploring cognitive factors like social status (Dalmaso et al., 2011; Jones et al., 2010) and judgments related to agents (Langton, 2009), have shown their impact on spatial attention bias during social interactions by regulating GCE. Given the pivotal role of GCE in human interaction (Frischen et al., 2007), identifying factors influencing GCE is crucial, aligning with the goal of our investigation.

Facial features serve as crucial social cues, typically leading individuals to draw inferences about others based on their facial appearance (Gawronski & Quinn, 2013; Ma et al., 2016; Rule et al., 2009; Todorov et al., 2005, 2015; Willis & Todorov, 2006). Trustworthiness emerges as a primary facet in the initial formation of facial impressions, as underscored in research by Oosterhof and Todorov (2008). Importantly, numerous studies have indicated that individuals can spontaneously form stable impressions of trustworthiness solely from facial characteristics within a brief time frame (Jones et al., 2021; Klapper et al., 2016; Marzi et al., 2014; Todorov et al., 2008a).

When defining trust, Hale et al. (2018) emphasize its multifaceted nature, with diverse definitions proposed by various scholars (Ashraf et al., 2006; Ben-Ner & Halldorsson, 2010). Broadly speaking, an individual's trust level can be classified into two categories. It may manifest as a stable personal trait, commonly referred to as "generalized trust," which signifies a general inclination to trust (Couch & Jones, 1997; Freitag & Traunmüller, 2009). Alternatively, it can be a more targeted response, which is based on the perceived attributes of a specific individual and fosters a sense of positive anticipation within peer groups, as articulated by Butler and Cantrell (1984) and further supported by McKnight et al. (1998). In our investigation, we adopt this latter perspective to explore the role of trustworthiness in shaping the GCE.

Gibson's (1979) ecological approach to social perception emphasizes the dynamic process of detecting and adjusting to environmental features, driving human adaptability in diverse situational contexts. Within this framework, spontaneous inferences of trustworthiness from facial expressions emerge as significant social cues that guide human behavior, influencing tendencies toward approach or avoidance (Todorov et al., 2008b). These inferences are integral in informing individuals on adapting effectively to their

social and environmental surroundings. Recognizing the social relevance of facial features and the consistent human inclination to form judgments of trustworthiness based on facial cues, it is reasonable to hypothesize that the extraction of trustworthiness information from facial stimuli plays a crucial role in directing attention.

While some previous research has attempted to address the question of whether trustworthiness can modulate the GCE, the findings have exhibited some inconsistencies. We present a selection of studies demonstrate both non-supportive and supportive perspectives. King et al. (2011) used vignettes to prime face trustworthiness, finding no significant moderation effect on the GCE during object categorization tasks. Similarly, Strachan et al. (2017) utilized pre-existing face stimuli varying in trustworthiness and required participants to complete object categorization, failing to demonstrate a modulation of trustworthiness on the GCE. Conversely, Süßenbach and Schönbrodt (2014) primed trustworthiness using brief personality descriptions, revealing a pronounced GCE with trustworthy faces during object localization task. Likewise, Jessen and Grossmann (2020) showed increased attention allocation to objects associated with trustworthy faces using event-related potentials (ERPs) and face stimuli with varying trustworthiness levels from the Oosterhof and Todorov (2008) database. In existing research on whether trustworthiness influences GCE, disparities arise from varied manipulation techniques and diverse tasks, leading to contradictory outcomes even in seemingly similar experimental designs (see King et al., 2011; Süßenbach & Schönbrodt, 2014). Considering these divergent perspectives, we intend to put forth two plausible hypotheses that may underlie the divergent result patterns. Subsequently, we aim to explore an enhanced methodology to discern which hypothesis aligns more accurately with empirical observations.

The reflexive hypothesis postulates the automatic nature of gaze following. Previous investigations into GCE have demonstrated that the direction of another person's gaze spontaneously induces a shift in the attention of observers (for a review, see Langton et al., 2000). For instance, despite informing participants of target likelihood on the opposite side, experiments by Driver et al. (1999) demonstrated faster discrimination on the gazed side, highlighting the automatic nature. Additionally, Law et al. (2010) found that GCE exhibited minimal dependence on cognitive resources. Assessing the impact of working memory load by having participants hold one or five digits during gaze trials, they observed consistent GCE magnitudes across both low and high memory load conditions. Even without conscious awareness, Sato et al. (2007) found that subliminally presented gaze cues still elicited shorter response times, supporting the automaticity of GCE. In light of this body of evidence, *the reflexive hypothesis* posits that trustworthiness

information associated with faces may not significantly influence gaze following since the cues inherently prompt an automatic shift in attention.

Contrary to *the reflexive hypothesis*, *the flexible hypothesis* suggests that GCE is not solely automatic but can be influenced by various factors. These modulators encompass factors such as social competence (Capozzi et al., 2016), communicative intent (Myllyneva & Hietanen, 2015; Senju & Csibra, 2008), knowledge about others' mental states (Teufel et al., 2010a), and the interplay between the social aspects stemming from eye contact and the strategic control related to cue validity (Kompatsiari et al., 2022). For instance, Teufel et al. (2010a) found that observers were more likely to follow gaze cues when they perceived the cue provider as having visual capability (transparent goggles) compared to situations where they believed the person had a lack of visual capability (opaque goggles). However, Cole et al. (2015), in a subsequent study using a gaze cue agent in a cardboard box, manipulating the agent's object perception ability, did not find support for the obligatory influence of attributed mental states on gaze cues, contrary to Teufel et al. (2010a). In multi-agent contexts, Capozzi et al. (2021) found a flexible interplay between cue numerosity and quality of social information, highlighting the adaptable nature of attention capture by gaze cues. In our research, we propose that if *the flexible hypothesis* holds, trustworthy faces may elicit a more pronounced GCE compared to untrustworthy facial cues.

Although there have been several studies examining the impact of trustworthiness on the GCE, there are opportunities for further refinement for several reasons. First, while prior studies have indirectly manipulated trustworthiness through methods such as priming neutral faces with brief personality descriptions (King et al., 2011; Süßenbach & Schönbrodt, 2014), these manipulations may introduce cognitive and affective elements. Our aim is to focus on understanding the fundamental perceptual processes to gain deeper insights into the intrinsic impact of trustworthiness. However, researchers have encountered inconsistencies in revealing trustworthiness-related features. For instance, Todorov et al. (2008a) identified the brow ridge, cheekbone, chin, and nose sellion as strongly correlated with trustworthiness judgments. Besides, Dotsch and Todorov (2012) observed that information in the mouth, eye, eyebrow, and hair regions also played a significant role in shaping judgments of trustworthiness. To ensure we could manipulate trustworthiness changes while keeping other dimensions unaffected, we employed a data-driven approach inspired by Oosterhof and Todorov's (2008) model for social perception of faces. They adopted principal components analysis (PCA) to identify key features, allowing representation of holistic sets of feature changes without prior assumptions about the significance of specific facial components.

Second, we rigorously exclude the confounding variables unrelated to trustworthiness perception. In experiments 2 and 3, we utilized stimuli from Oosterhof and Todorov (2008)'s database. Their framework posits that trustworthiness and dominance contribute to the 2D structure of face evaluation, aligning with established social perception models (Fiske et al., 2007; Wiggins, 1979). Their methodology enabled changes in trustworthiness while maintaining other trait judgments within the same face identity. Morphing facial stimuli by 3 standard deviations in trustworthiness (experiment 2) and setting them to baseline trustworthiness (experiment 3) enables a focused examination of trustworthiness in the GCE, minimizing interference from confounding variables related to face judgments. This stands in contrast to studies manipulating trustworthiness perception across different face identities. The absence of a comparison of the GCE effect within the same identity with only trustworthiness changed may introduce potential confounding variables, such as the influence of other dimensions like dominance in the 2D structure of face evaluation. This oversight could contribute to the mixed findings regarding the influence of trustworthiness on the GCE.

Third, our study diverges from previous research, which predominantly favored employing object categorization tasks to gauge participants' responses (e.g., Bayliss & Tipper, 2006; King et al., 2011; Strachan et al., 2017). Object categorization tasks involve higher-level cognitive functions associated with grouping and recognizing visual stimuli within meaningful categories. In contrast, our approach centered around an orientation discrimination task. This task specifically targets basic visual perception abilities, focusing on fundamental processes rather than complex object identification. As outlined in our first point, our primary objective is to delve into the core perceptual mechanisms to gain a deeper understanding of the intrinsic impact of trustworthiness.

In sum, the primary objective of the present study was to investigate whether trustworthiness perceived from facial visual appearance affects GCE. Specifically, we employed facial stimuli that had been pre-rated for trustworthiness by a separate group (categorized as trustworthy or untrustworthy). These faces served as cues with manipulated gaze directions, and participants were required to promptly indicate the orientation of the Gabor probe. In Experiment 1, we utilized facial stimuli from the Klapper et al. (2016) study. Our aim was to investigate whether the degree of trustworthiness in these facial stimuli would impact the magnitude of the GCE. For Experiment 2, we sought to replicate the findings of Experiment 1 using facial stimuli sourced from the Oosterhof and Todorov (2008) database. In Experiment 3a and 3b, we standardized the trustworthiness information of the facial stimuli employed in Experiment 2 to establish a baseline for a control experiment. This approach enabled

us to examine whether any effects observed in Experiment 1 and 2 were influenced by potential confounding variables, apart from trustworthiness. If the flexible hypothesis were true, it would imply that trustworthiness information from faces could indeed modulate the GCE in Experiments 1 and 2. Furthermore, when we set the trustworthiness information to baseline in Experiments 3a and 3b, the modulation effect of trustworthiness on GCE may disappear. Conversely, if the reflexive hypothesis were accurate, we would expect no modulation effect of trustworthiness in any of the experiments. The raw data for all these studies can be accessed on the Open Science Framework website (<https://osf.io/y6xhf>).

Experiment 1: the trustworthiness of faces can modulate the GCE

Methods

Participants

Thirty-six students (19 female, age range: 18–26 years, mean age: 20.17 years) were recruited from Shandong Normal University. All participants reported normal or corrected to normal vision. Participants were naïve to the purpose of the study and provided written consent before participation. The study was approved by the Research Ethics Board of Shandong Normal University. Each participant was compensated with a monetary reward of 15 CNY for their participation in this experiment.

To ensure adequate power, the sample size was determined by a power analysis based on the predicted effect size using PANGEA (Westfall, 2015). Based on the results of previous studies (e.g., Süßenbach & Schönbrodt, 2014), we predicted a large effect size ($d=0.7$, according to Cohen, 1988) for the interaction effect in our experimental design. With 90% power at a 0.05 significance level, the suggested sample size was approximately 22 individuals. This sample size determination method was applied consistently across all experiments.

Initially, we recruited 24 participants for each of experiments 1 and 2. However, given the inconsistency observed in previous literature on this topic, an additional 12 participants were recruited for each experiment to enhance the stability and robustness of our results, resulting in a final sample size of 36. It's noteworthy that the overall findings from the total sample of 36 participants are in line with those from the initial sample of 24 participants. Details regarding both the original and newly recruited samples are provided in Appendix A. Furthermore, to evaluate the potential impact of sample recruitment status (original vs. newly sampled) on our findings, we included sample status as a between-subject factor and found no significant influence (see Appendix B).

Design

For Experiment 1, we used a 2 (face trustworthiness: trustworthy and untrustworthy) \times 2 (gaze-cue type: valid and invalid) within-subject design. On one-half of the trials, the faces were trustworthy, and on the other, untrustworthy. The gaze direction (left or right) was consistent with the position (left or right of the fixation) of the Gabor probe on one-half of the trials (valid trials) and inconsistent on the other half of the trials (invalid trials). Participants completed 256 trials, which were divided into two blocks. For each block, the four conditions mentioned above were fully randomized, and each face was repeated eight times.

Stimuli

We referenced the Klapper et al. (2016) database, which consisted of eight trustworthy looking faces and eight untrustworthy looking faces ($12.12^\circ \times 14.45^\circ$; see Fig. 1). We used all the faces from this database in our experiment. The facial materials were generated using the FaceGen software development kit (Oosterhof & Todorov, 2008). Initially, eight male speakers' standard average faces were morphed two standard deviations towards male characteristics. Then, each face was further altered by morphing it six standard deviations on a random dimension orthogonal to known social dimensions, creating distinct individual features. Individual overlay textures were applied to enhance realism, adding details like skin irregularities. Additionally, each face was given a unique haircut from real face images in the Radboud Face Database. Importantly, trustworthy and untrustworthy versions of each of these eight faces were created by morphing them 2.5 standard deviations toward either trustworthiness or untrustworthiness. Besides, Oosterhof and Todorov (2008) highlights the similarities in trait evaluations between natural and computer-generated faces. The correlations between trait ratings, as well as the results of



Fig. 1 Examples of *trustworthy* (left) and *untrustworthy* faces (right)

PCA analyses, were found to be consistent, underlining the validity of using computer-generated faces as experimental materials. Before the experiment, we recruited two groups of 30 Chinese participants (15 females, age range: 19–25 years, mean age: 21.0 years) and asked them to rate the degree of trustworthiness for the adopted face stimuli on a seven-point scale (1 = very untrustworthy, 7 = very trustworthy). Results of the trustworthiness evaluation showed that ratings for face stimuli used here were significantly higher for trustworthy ($M = 4.47$, $SD = 0.911$) compared with untrustworthy faces ($M = 2.91$, $SD = 0.912$), $t(29) = 8.01$, $p < 0.001$, Cohen's $d = 1.46$. The rating results confirmed that the trustworthiness manipulation was both successful and robust for Chinese participants.

We manipulated each face to produce left-gaze and right-gaze cues by shifting the pupil and iris area of each eye into the left and right corner of each eye, using GIMP software (version 2.8.16, GIMP Development Team, 2016). The average degree of gaze change is 0.17° . Ultimately, we adopted 24 trustworthy looking faces (i.e., eight different faces in straight-gaze/ left-gaze/ right-gaze version each) and 24 untrustworthy looking faces (i.e., eight different faces in three version each). Dynamic eye cues were used in the current research for they could extend the work on attention to a more realistic situation (Hermens & Walker, 2012; Kuhn & Tipples, 2011).

For the current experiment, two Gabor patches ($1.5^\circ \times 1.5^\circ$, 3° clockwise or counter-clockwise rotation, and 4.0 cycles per degree) were used as probes. A 200-ms, 2000-Hz beep was used as the error feedback.

Apparatus

Visual stimuli were generated on a PC using PsychoPy (Peirce, 2007), and presented on a linearized CRT monitor (21" Sun GDM-5510; resolution: 1024×768 pixels; refresh rate: 100 Hz). Participants viewed the screen from a distance of

70 cm in a dark room. Stimuli were presented against a uniform gray background at a mean luminance of 40 cd/m^2 .

Procedure

As illustrated in Fig. 2, each trial began with a black fixation cross (500 ms), followed by a face with dynamic gaze (500-ms straight gaze + 500-ms left or right gaze) and then the 100-ms ISI display. Next, a Gabor patch was presented on the left (or right) side of the fixation until the participant's response or for a maximum of 1500 ms. The Gabor patch was presented slightly tilted either clockwise or counter-clockwise. Participants were instructed to promptly press one of two keys to indicate their perceived orientation of the Gabor patch, regardless of its side of presentation. The tilted directions of Gabor patches were counterbalanced across different conditions. The inter-trial interval was 800–1200 ms. At the beginning of the experiment, participants were explicitly told that neither of the gaze directions was predictive of probe location (Petrican et al., 2012) and the primary objective was to respond swiftly while minimizing errors.

Statistical analysis

The statistical analysis was carried out using JASP (JASP Team, 2024), employing repeated measures analysis of variance (ANOVA). Initially, ANOVA was applied to the mean accuracy rate with within-subject factors: face trustworthiness (trustworthy and untrustworthy) and gaze-cue type (valid and invalid). Previous studies exploring influence of trustworthiness on GCE usually measured reaction times (RTs) as an index (King et al., 2011; Petrican et al., 2012; Süßenbach & Schönbrodt, 2014). Therefore, in our research, we referenced the previous literature and made RTs the main measure of interest. We also reported accuracy results for the sake of completeness.

For the analysis of RTs, to enhance the robustness of our analysis, we adhered to the methodology outlined by

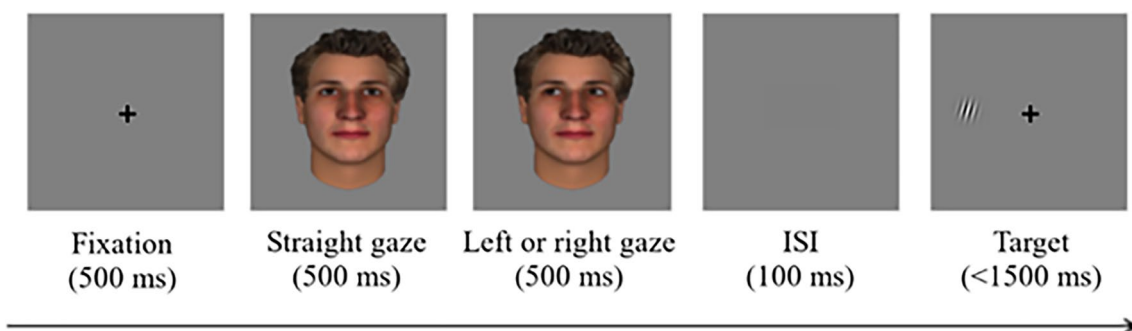


Fig. 2 Procedure for Experiment 1. Each trial began with a black fixation cross (500 ms), followed by a face with dynamic gaze (500-ms straight gaze + 500-ms left or right gaze) and then the 100-ms ISI display.

Next, a Gabor probe was presented on the left (or right) side of the fixation until the participant pressed either the left or right arrow key on the keyboard or for a maximum of 1500 ms

Petrican et al. (2012), which involves excluding trials with RTs deviating more than three standard deviations from the participant mean. This precautionary step aimed to minimize the impact of extreme outlier values on our results and ensure a more reliable dataset. Subsequently, a two-way repeated-measures ANOVA was conducted for the remaining correct RTs, employing the same within-subject factors as in the accuracy analysis. Our exclusion criterion resulted in the exclusion of 393, 363, 323, and 334 trials, constituting 4.26%, 3.94%, 5.26%, and 3.62% of the total number of trials in Experiment 1, 2, 3a, and 3b, respectively. For further scrutiny of the GCE, which signifies the distinction between RTs in the valid condition and RTs in the invalid condition, post hoc multiple comparisons were conducted using the Holm's correction method (Holm, 1979) if the interaction effect reached significance. It's worth noting that while the post hoc test encompassed all possible comparisons, only the contrasts pertinent to our research goals were reported. These contrasts specifically focused on the comparison between RTs of valid and invalid conditions for trustworthy and untrustworthy stimuli, respectively.

To provide additional support for our findings, we computed Bayes Factors using JASP with default prior width for all the analyses we performed above. As outlined by Wagenmakers et al. (2018), the Bayes Factor is denoted as BF_{10} (and its reciprocal $BF_{01} = 1/BF_{10}$), quantifying the strength of evidence that the data offer for the alternative hypothesis in comparison to the null hypothesis. We interpret Bayes Factors as follows: values below 3 indicate anecdotal

evidence, those falling within the range of 3 to 10 suggest moderate evidence, values between 10 and 30 signify strong evidence, the range of 30 to 100 represents very strong evidence, and values exceeding 100 are indicative of decisive evidence (Gray et al., 2018; Jeffreys, 1961).

Results

In Experiment 1, we tested whether the trustworthiness of faces can affect the GCE using stimuli sourced from the Klapper et al. (2016) database. Descriptive data for Experiment 1 (and the other experiments) were presented in Table 1.

RTs

The analysis revealed significant main effects for gaze-cue type [valid vs. invalid: 647 vs. 660 ms, $MSE = 477.528$, $F(1, 35) = 12.525$, $p = 0.001$, $\eta_p^2 = 0.264$, $BF_{10} = 31.568$] and an interaction between trustworthiness and gaze-cue type [$MSE = 160.332$, $F(1, 35) = 4.297$, $p = 0.046$, $\eta_p^2 = 0.109$, $BF_{10} = 2.186$]. Notably, there was no significant main effect of trustworthiness [trustworthy vs. untrustworthy: 653 vs. 653 ms, $MSE = 153.043$, $F(1, 35) = 0.014$, $p = 0.906$, $\eta_p^2 < 0.001$, $BF_{01} = 4.630$; see Fig. 3].

We conducted post hoc multiple comparisons using Holm's correction method. For trustworthy faces, a GCE

Table 1 Descriptive results of all the experiments

			Trustworthy-valid	Trustworthy-invalid	Untrustworthy-valid	Untrustworthy-invalid
Experiment 1	RT	M	644.82	662.08	648.95	657.46
		SD	69.90	72.83	71.40	69.79
	ACC	M	0.98	0.97	0.97	0.97
		SD	0.03	0.02	0.03	0.03
Experiment 2	RT	M	627.66	639.64	636.96	631.16
		SD	57.61	61.86	58.30	61.62
	ACC	M	0.97	0.97	0.98	0.98
		SD	0.05	0.04	0.05	0.03
Experiment 3a	RT	M	634.99	644.69	637.35	647.40
		SD	69.16	70.10	71.48	72.27
	ACC	M	0.96	0.96	0.96	0.96
		SD	0.09	0.05	0.05	0.04
Experiment 3b	RT	M	685.52	701.18	685.81	707.68
		SD	99.45	102.25	100.16	99.88
	ACC	M	0.98	0.98	0.97	0.98
		SD	0.03	0.03	0.03	0.03

Descriptive data for Experiments 1, 2, 3 is presented, encompassing RTs and accuracy rate (ACC) statistics. The data includes the mean (M) and standard deviation (SD) values of RTs (ms) and ACC for each experimental condition

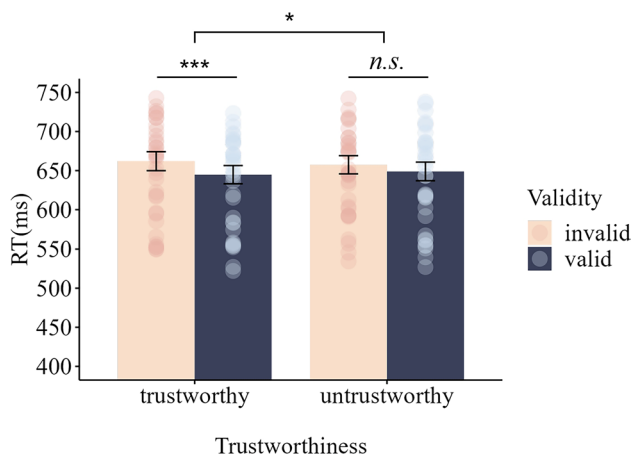


Fig. 3 Results of Experiment 1. Mean RTs were calculated for facial cues categorized as trustworthy and untrustworthy in both valid and invalid trial conditions. Error bars represent standard errors. Significance levels are denoted as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and “n.s.” indicating no significant difference

was detected, with a mean GCE of 17 ms, $t(35) = 4.102$, $p < 0.001$, Cohen’s $d = 0.243$. However, there was no significant GCE observed for untrustworthy faces, with a mean GCE of 9 ms, $t(35) = 2.023$, $p = 0.144$, Cohen’s $d = 0.120$. These findings indicate that the perceived trustworthiness of facial appearance has a discernible impact on the GCE.

Accuracy

The overall accuracy was 97%. A repeated-measures ANOVA with the trustworthiness and gaze-cue type as the within-subjects factors revealed that neither the main effects nor the interaction effect was significant (trustworthiness: $F(1, 37) = 0.243$, $p = 0.625$, $\eta_p^2 = 0.007$, $BF_{01} = 4.310$; gaze-cue type: $F(1, 37) = 2.476$, $p = 0.125$, $\eta_p^2 = 0.066$, $BF_{01} = 2.137$; interaction effect: $F(1, 37) = 0.508$, $p = 0.481$, $\eta_p^2 = 0.014$, $BF_{01} = 2.976$).

Discussion

In Experiment 1, we utilized facial stimuli from the database developed by Klapper et al. (2016), varying along the dimensions of trustworthiness and untrustworthiness. Our findings revealed that participants exhibited faster RTs in response to the orientation of the Gabor patch when the gaze cue was valid. However, there were no significant differences in RTs between the trustworthy and untrustworthy conditions. Importantly, when the facial stimuli were perceived as trustworthy, they elicited a significantly stronger GCE compared to situations where the stimuli were deemed untrustworthy. Such findings support the *flexible hypothesis*, which posits

that GCE could be modulated by trustworthiness but not reflexive processes.

Notably, our observed results pattern can be interpreted such that trustworthy faces facilitated the GCE, or untrustworthy faces impeded it. Considering the previous finding suggesting that gaze cues have the capability to direct attention independently, even without explicit trustworthiness manipulations (for an overview, refer to Frischen et al., 2007), we may prefer possibility that, rather than trustworthiness enhancing the GCE, untrustworthiness may have a hindering effect on this attentional phenomenon. While the reported p -value for the interaction effect achieved statistical significance, the Bayes Factor did not surpass the threshold of 3, suggesting only anecdotal evidence for the interaction effect of trustworthiness and GCE. Consequently, a replication of Experiment 1 was conducted to confirm the robustness of our results. This led to the initiation of Experiment 2, serving as a replication study of Experiment 1.

Experiment 2: the trustworthiness of faces can modulate the GCE (Replication study of experiment 1)

Methods

In Experiment 2, we adopted another face stimuli set from the Oosterhof and Todorov (2008) database and aimed to confirm the robustness of the results in Experiment 1. The facial stimuli were generated via computer using FaceGen Modeller 3.2 software (Singular Inversions, 2007, Toronto, Canada) and were manipulated to represent different levels of trustworthiness based on the model outlined by Oosterhof and Todorov (2008). Furthermore, the facial trustworthiness information in this database could be adjusted to baseline for the control experiment later.

Participants

Thirty-six students (12 female, age range: 18–20 years, mean age: 19.2 years) were recruited from Sun Yat-sen University. All participants reported normal or corrected to normal vision. Participants were naïve to the purpose of the study and provided written consent before participation. The study was approved by the Research Ethics Board of Sun Yat-sen University.

Stimuli

Eight trustworthy looking faces and eight untrustworthy looking faces ($12.12^\circ \times 14.45^\circ$) were randomly selected from the Oosterhof and Todorov (2008) database (see Fig. 4).



Fig. 4 Examples of trustworthy (left) and untrustworthy faces (right)

As in Experiment 1, we recruited another group of 30 Chinese participants (15 females, age range: 19–25 years, mean age: 20.8 years) before the experiment and asked them to rate the degree of trustworthiness for face stimuli on a seven-point scale (1 = very untrustworthy, 7 = very trustworthy). Results of the trustworthiness evaluation showed that ratings of trustworthy faces ($M = 4.92$, $SD = 0.891$) were significantly higher than untrustworthy faces ($M = 2.91$, $SD = 0.912$), $t(29) = 9.12$, $p < 0.001$, Cohen's $d = 2.40$. The results confirmed that these facial stimuli worked well in Chinese samples.

The design, apparatus, procedure, and statistical analysis were the same as in Experiment 1.

Results

RTs

The results revealed no significant main effects of trustworthiness and gaze-cue type [trustworthiness (trustworthy vs. untrustworthy): 634 vs. 634 ms, $MSE = 115.396$, $F(1, 35) = 0.054$, $p = 0.818$, $\eta_p^2 = 0.002$, $BF_{01} = 4.695$; gaze-cue type (valid vs. invalid): 632 vs. 635 ms, $MSE = 225.332$, $F(1, 35) = 1.528$, $p = 0.225$, $\eta_p^2 = 0.431$, $BF_{01} = 1.815$]. However, there was a significant interaction between trustworthiness and gaze-cue type [$MSE = 107.669$, $F(1, 35) = 26.420$, $p < 0.001$, $\eta_p^2 = 0.414$, $BF_{10} = 3.5863 \times 10^4$; Fig. 5].

Through post hoc tests, we identified the presence of GCE from trustworthy faces [12 ms, $t(35) = 3.939$, $p = 0.001$, Cohen's $d = 0.200$]. However, there was a reversed GCE from untrustworthy faces [-6 ms, $t(35) = -1.906$, $p = 0.184$, Cohen's $d = -0.097$]. Consistent with findings in Experiment 1, these results showed

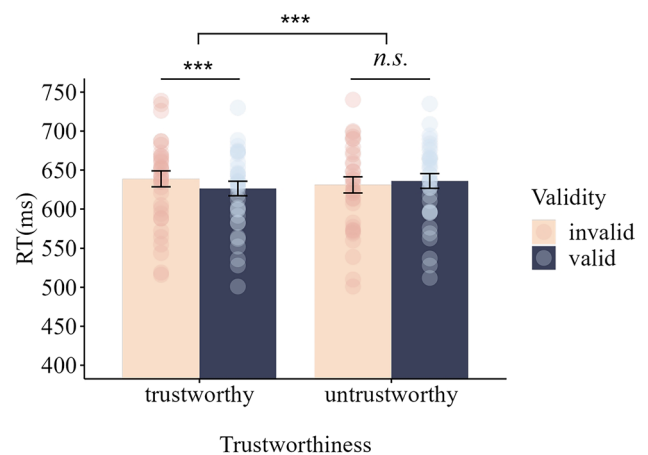


Fig. 5 Results of Experiment 2

that the trustworthiness impression from facial appearance affected the GCE.

Accuracy

The overall accuracy rate reached 97%. A repeated-measures ANOVA indicated that the main effect of gaze-cue type nor the interaction effect achieved statistical significance [$F(1, 35) = 0.002$, $p = 0.967$, $\eta_p^2 < 0.001$, $BF_{01} = 2.994$; interaction effect: $F(1, 35) = 1.688$, $p = 0.202$, $\eta_p^2 = 0.046$, $BF_{01} = 1.842$]. The main effect of trustworthiness exhibited statistical significance [$F(1, 35) = 5.924$, $p = 0.020$, $\eta_p^2 = 0.145$, $BF_{10} = 1.708$], with a 97.05% accuracy rate in the trustworthy condition and a 97.68% accuracy rate in the untrustworthy condition. However, given the generally high accuracy rate, the importance of this effect may be limited.

Discussion

In Experiment 2, we utilized facial stimuli from the database created by Oosterhof and Todorov (2008) and aimed to replicate the findings obtained in Experiment 1. Despite the factors of validity and trustworthiness not individually yielding differences in RTs, it is worth highlighting that when the facial stimuli were perceived as trustworthy, they elicited a significant GCE. Conversely, for facial stimuli perceived as untrustworthy, we observed a subtle reversal in the GCE tendency, indicating even shorter RTs when the gaze cue was invalid. This, to some extent, can provide support for the proposal we posited in the discussion section of Experiment 1, namely that untrustworthiness impedes the GCE. Experiment 2 also lent support for the *flexible hypothesis*.

In Experiment 2, the Bayes Factor consistently corroborated all the findings revealed by p values, with particular emphasis on the BF_{10} of the interaction effect, providing

decisive evidence in accordance with Jeffreys (1961). This successful replication of the modulation effect of trustworthiness on the GCE in Experiment 2 further underpins the robustness of the findings obtained in Experiment 1.

We observed conflicting results between accuracy and RTs in the main effects of trustworthiness, despite having mentioned that we primarily focus on RTs as the main index. While accuracy reached statistical significance, the value of BF_{10} suggested only anecdotal evidence for this effect. Additionally, such significant main effects of trustworthiness on accuracy were not replicated in any other studies in our research. Moreover, the accuracy rates for the trustworthy and untrustworthy conditions were 97.05% and 97.68%, respectively, both indicating a relatively high standard of accuracy. Thus, we hypothesize that the practical significance of this difference may be trivial.

Experiment 3a/3b: the face stimuli without trustworthiness cannot modulate the GCE

In the Oosterhof and Todorov (2008) database, the trustworthiness level of facial stimuli can be modified on a scale ranging from -3 to +3, while other facial attributes remain unchanged. For instance, the same face may exhibit a high trustworthiness level (+3) or a low trustworthiness level (-3), as we employed in Experiment 2. In this context, a trustworthiness score of 0 serves as the baseline, indicating facial stimuli with a trustworthiness level set at zero. In Experiment 3, we reset the trustworthiness information of the faces utilized in Experiment 2 to the baseline level, effectively transforming them into baseline faces. This manipulation was undertaken to investigate the influence of trustworthiness on GCE, while keeping other low-level characteristics consistent. We hypothesized that, following this baseline manipulation, no significant differences would be observed in RTs between the trustworthy and untrustworthy faces employed in Experiment 2. To ensure the stability of our results, we replicated Experiment 3a in Experiment 3b with a larger sample size to assess the replicability of the findings.

Methods

Participants

In Experiment 3a, twenty-four participants (12 females, age range: 18–20 years, mean age: 18.8 years) were recruited, while in Experiment 3b, thirty-six participants (22 females, age range: 19–27 years, mean age: 21.3 years) were recruited from Sun Yat-sen University.

All participants reported normal or corrected to normal vision. Participants were naïve to the purpose of the study and provided written consent before participation. The study

was approved by the Research Ethics Board of Sun Yat-sen University.

Stimuli

The face stimuli used in Experiment 2 from Oosterhof and Todorov (2008) allowed us to manipulate the degree of trustworthiness. We still adopted the faces in Experiment 2, except that we set the trustworthiness information to baseline here (see Fig. 6).

As in previous experiments, a group of 30 Chinese participants (16 females, age range: 19–21 years, mean age: 20.5 years) were recruited before the experiment and asked to rate the degree of trustworthiness for face stimuli on a seven-point scale (1 = very untrustworthy, 7 = very trustworthy). Results of the trustworthiness evaluation showed that ratings of used-to be trustworthy faces ($M=4.33$, $SD=0.793$) were no difference to untrustworthy faces ($M=4.11$, $SD=0.849$), $t(29)=1.012$, $p=0.316$, Cohen's $d=0.261$. The results confirmed that the baseline setting manipulation of trustworthiness information was successful.

The design, apparatus, procedure, and statistical analysis were the same as in Experiment 1, except that we additionally performed the Bayesian ANOVA to provide support for the null hypothesis of interaction effect with JASP (Version 0.17.2).

Results of experiment 3a

RTs

The results revealed a significant main effect of gaze-cue type: [valid vs. invalid: 636 vs. 646 ms, $MSE=51.036$, $F(1,$



Fig. 6 Examples of setting trustworthy (left) and untrustworthy faces (right) to baseline faces (i.e., trustworthiness information set to zero)

23)=45.897, $p < 0.001$, $\eta_p^2 = 0.666$, $BF_{10} = 1.151 \times 10^4$]. However, the main effect of trustworthiness [trustworthy vs. untrustworthy: 640 vs. 642 ms, $MSE = 135.019$, $F(1, 23) = 1.139$, $p = 0.297$, $\eta_p^2 = 0.047$, $BF_{01} = 1.876$] and the interaction effect [$MSE = 41.156$, $F(1, 23) = 0.018$, $p = 0.896$, $\eta_p^2 = 0.001$, $BF_{01} = 3.663$; Fig. 7] were not significant.

Accuracy

The overall accuracy was 96%. A repeated-measures ANOVA revealed that neither the main effects nor the interaction effect was significant (trustworthiness: $F(1, 23) = 0.240$, $p = 0.629$, $\eta_p^2 = 0.010$, $BF_{01} = 3.460$; gaze-cue

type: $F(1, 23) = 0.175$, $p = 0.679$, $\eta_p^2 = 0.008$, $BF_{01} = 2.770$; interaction effect: $F(1, 23) < 0.001$, $p = 0.997$, $\eta_p^2 < 0.001$, $BF_{01} = 3.717$).

Results of experiment 3b

RTs

The results revealed a significant main effect of gaze-cue type, with valid cues (686 ms) leading to shorter RTs compared to invalid cues (704 ms) [$MSE = 341.704$, $F(1, 35) = 37.173$, $p < 0.001$, $\eta_p^2 = 0.515$, $BF_{10} = 2.351 \times 10^4$]. However, the main effect of trustworthiness did not reach significance [trustworthy vs. untrustworthy: 693 vs. 697 ms, $MSE = 258.054$, $F(1, 35) = 1.607$, $p = 0.213$, $\eta_p^2 = 0.044$, $BF_{01} = 2.604$], and there was no significant interaction effect [$MSE = 220.233$, $F(1, 35) = 1.575$, $p = 0.218$, $\eta_p^2 = 0.043$, $BF_{01} = 1.786$; Fig. 7].

Accuracy

The overall accuracy stood at 96%. A repeated-measures ANOVA indicated that neither the main effects nor the interaction effect was significant (trustworthiness: $F(1, 35) = 0.227$, $p = 0.637$, $\eta_p^2 = 0.006$, $BF_{01} = 3.831$; gaze-cue type: $F(1, 35) = 0.213$, $p = 0.648$, $\eta_p^2 = 0.006$, $BF_{01} = 4.505$; interaction effect: $F(1, 35) = 1.154$, $p = 0.290$, $\eta_p^2 = 0.032$, $BF_{01} = 1.901$).

Discussion

Experiment 3 serves as a control experiment aimed at detecting any null effects in the interaction, thus ensuring that the observed changes in GCE were specifically attributable to trustworthiness rather than other low-level facial features. Given the importance of ensuring that even small effect sizes were not overlooked, we chose to replicate Experiment 3a with an increased sample size of 36 participants in Experiment 3b, thus enhancing the validity of our preceding findings.

The results of Experiment 3a and 3b consistently exhibit a clear pattern. When the trustworthiness information was removed from the facial stimuli employed in Experiment 2, the interaction effect shifted from being statistically significant in Experiment 2 to becoming non-significant in Experiment 3. Additionally, we examined the absence of the interaction effect from a Bayesian perspective, aligning with the main objective of Experiment 3. The BF_{01} for the interaction effect in Experiment 3a was 3.663, indicating moderate evidence in favor of a null effect. Besides, in Experiment 3b, the BF_{01} for the interaction effect was 1.786, providing anecdotal evidence supporting the null hypothesis. The

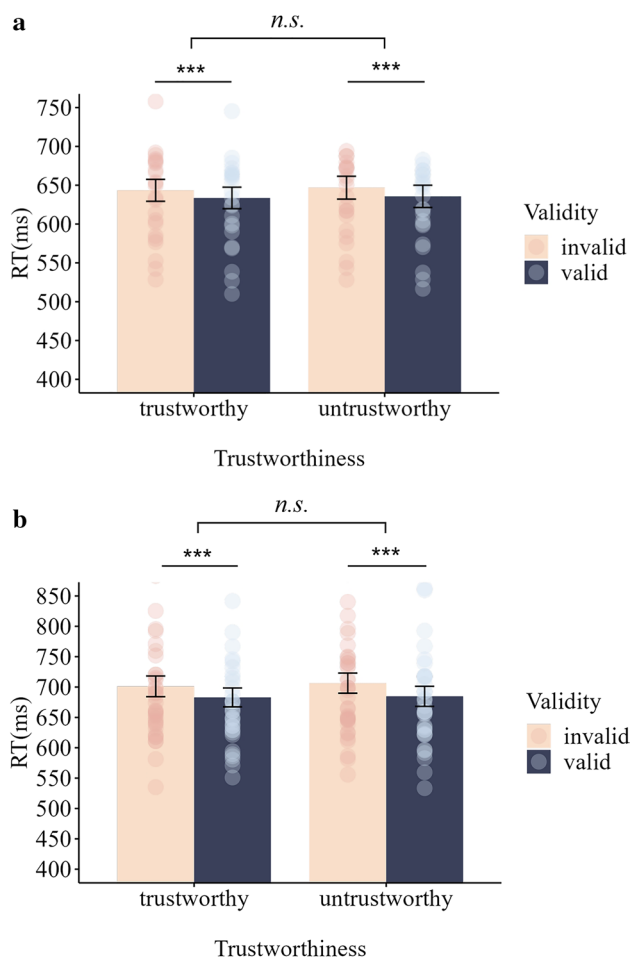


Fig. 7 Results of Experiment 3a/3b. Mean RTs were computed for facial cues that were previously categorized as trustworthy and untrustworthy in both the valid and invalid trial conditions of Experiment 3. It is important to note that in Experiment 3, the facial stimuli were intentionally adjusted to be neutral in the trustworthiness dimension. Therefore, there were no inherently trustworthy or untrustworthy faces in this experiment, but rather faces that had been rated as trustworthy or untrustworthy in Experiment 2 prior to the relevant features being altered. (a) RTs were calculated for the Experiment 3a (N=24). (b) RTs were calculated for Experiment 3b (N=36)

results suggested that there was a GCE effect in general. These findings provide support for the conclusion that trustworthiness is indeed the factor modulating the GCE, providing evidence in favor of the *flexible hypothesis*. This outcome aligns with our hypothesis that the GCE is present even for facial stimuli rated as neutral in trustworthiness, and that the modulation effect of trustworthiness primarily manifests when untrustworthiness hinders the GCE, rather than trustworthiness enhancing it.

General discussion

The current study has revealed that trust plays a crucial role in modulating the GCE, and this observed influence cannot be ascribed to trustworthiness-unrelated features of the facial stimuli. In Experiment 1, the GCE was evident in the trustworthy condition, whereas no such effect was observed in the untrustworthy condition. Building on this, in Experiment 2, we replicated the modulation effect of trustworthiness on the GCE using a new stimulus set sourced from Oosterhof and Todorov (2008). Importantly, this stimulus set allowed for the adjustment of trustworthiness information to a baseline level. In Experiment 3a, after resetting the trustworthiness information to baseline, we found no significant difference in the magnitude of the GCE between the trustworthy and untrustworthy faces utilized in Experiment 2. Furthermore, to enhance the robustness and validity of our findings in Experiment 3a and increasing statistical power, we involved an additional group of 36 students in Experiment 3b to replicate the results, successfully confirming the outcomes observed in Experiment 3a. These results collectively indicate that the effects observed in Experiment 1 and Experiment 2 were not confounded by trustworthiness-unrelated characteristics of the face stimuli. Taken together, our findings support the flexible hypothesis that attentional orienting triggered by gaze following is subject to modulation by top-down factors such as trustworthiness.

Regarding the modulation effect of trustworthiness on GCE, our study unravels a nuanced interpretation. Existing research has consistently shown that gaze cues, in isolation, possess the capacity to direct attention, even in the absence of explicit trustworthiness manipulations (for a review, see Frischen et al., 2007). Contrary to the facilitation by trustworthy faces, our findings suggest that untrustworthy faces impede GCE. Notably, in Experiment 3, we observed significant GCE for neutral faces with trustworthiness set to baseline. Conversely, in Experiment 2, no significant GCE was noted for untrustworthy faces, and there was even a slight reversed GCE, although this did not reach statistical significance. This phenomenon could be elucidated by Todorov et al., (2008a, 2008b) proposition that judgments of trustworthiness may play a crucial role in social interactions, precisely in determining whether to approach or avoid a stranger, particularly in the absence of clear emotional cues

indicating the intentions of the other person. Accordingly, in Experiment 2, when untrustworthy faces directed their gaze in a specific direction, individuals may have been predisposed to refrain from following their gaze and instead sought to redirect their attention elsewhere. This inclination could account for the observed quicker responses in invalid trials. The absence of a significant GCE in the untrustworthy condition in Experiment 1 further reinforces the notion that untrustworthy faces hinder GCE. To account for the discrepancy between the nonsignificant tendency of GCE in Experiment 1 and the presence of a non-significant reversed GCE in Experiment 2, we consider the nature of the stimuli used. In Experiment 1, we relied on a database established by Klapper et al. (2016), which involved morphing trustworthy and untrustworthy versions of each of the eight faces by shifting them 2.5 standard deviations toward trustworthiness or untrustworthiness, as per the model of Oosterhof and Todorov (2008). In contrast, for Experiment 2, we selected the most trustworthy and untrustworthy faces from the database of Oosterhof and Todorov (2008), morphed 3 standard deviations in the respective trustworthiness direction. Consequently, the stimuli used in Experiment 2 may have exhibited a higher degree of untrustworthiness compared to those in Experiment 1, thus potentially more effectively diverting participants' attention to another direction. Nevertheless, this is just one plausible speculation that should be further investigated in future research.

The present study contributes to the realm of flexible interpretation. Our findings substantiate the support for flexible interpretation by demonstrating that top-down influences from agents, encompassing intricate social factors such as trustworthiness, effectively modulate GCE. This extension of prior research brings to light the dynamic nature of perception in gaze following, challenging the notion of a rigid or purely reflexive human attention orientation system. Instead, we assert that this system demonstrates heightened flexibility and adaptability across diverse situational contexts, consistent with earlier research findings (Teufel et al., 2010b). Our ability to selectively attend to stimuli in the environment is indicative of a prioritized processing mechanism for utilitarian stimuli, enabling the conservation of resources, facilitation of social interactions, and enhanced efficiency in goal attainment. However, this attentional allocation is not fixed; rather, it dynamically adjusts based on perceived social factors like trustworthiness. Elucidated in this study, we propose that evaluating someone's trustworthiness based on facial appearance serves as a valuable means for individuals to efficiently extract utilitarian information when other cues are lacking. Notably, our study showcases a swift redirection of attention away from untrustworthy faces, indicative of a strategic response to conserve resources in the face of potential deception.

Considering the significance of GCE in a social context, one might question whether GCE is inherently flexible. In a recent publication (Zohary et al., 2022), researchers explored whether people can automatically develop GCE when gaze cues

become available only in late childhood. Results indicated that late-treated cataract patients failed to understand and follow eye gaze direction, suggesting that individuals rely on an unsupervised gaze-learning process during infancy to shape GCE. However, it remains unclear whether the modulation effect of social factors observed in our research is an outcome of acquired learning during infancy. Future research endeavors could further illuminate the link between primary characteristics and the more intricate top-down processes of social factor modulation identified in our study.

In our paradigm, there might be potential interference factors that require further explanation. First, some may question the use of the orientation discrimination task, as it may involve a second level of congruency where the direction of the Gabor patch tilt aligns with the gaze cue. However, it's worth noting that the adoption of the Gabor orientation discrimination paradigm has previously been validated in Shi et al. (2010), where it was employed to investigate the correlation between the walking direction of biological motion and attentional orienting, thus partially affirming its validity. Moreover, participants were instructed to discern whether the Gabor patch was tilted clockwise or counter-clockwise and respond promptly by pressing one of two keys accordingly. The directions of tilt for the Gabor patches were counterbalanced across different conditions, thereby reducing the likelihood of alignment between the tilt direction of the Gabor patch and the gaze cue. Second, examining the stimuli, it is evident that eye size is often altered across identities when manipulating trustworthiness. Systematic differences in eye size across conditions, typically with smaller eyes for less trustworthy conditions, may suggest that eye gaze information is less salient for smaller eyes, potentially

weakening the GCE. However, studies such as Bayless et al. (2011) demonstrated that emotional expressions (e.g., fear), rather than eye size, account for GCE. When maintaining constant eye size but altering the facial orientation by using face inversion or presenting only the eye regions, the GCE was diminished or even eliminated. Additional research, exemplified by the study conducted by Carlson and Aday (2018), supports the notion that eye size alone cannot solely drive attention guidance. Therefore, future studies manipulating trust information while maintaining constant eye size to investigate its impact on GCE would enhance the credibility of results. Moreover, our use of computer-generated faces as stimuli in the current study encourages future research to incorporate real faces to extend the validity of findings.

In summary, our study highlights that trustworthiness significantly modulates the GCE, with indications suggesting that untrustworthy faces may hinder the GCE. These results challenge the conventional view of GCE as a purely reflexive attentional shift, proposing instead that attentional orienting operates flexibly, supporting top-down processing of the GCE. This provides empirical support for the flexible hypothesis, underscoring the dynamic nature of attentional mechanisms in the context of social cues.

Appendix

See Tables 2, 3 and 4

Table 2 Descriptive results

			Trustworthy-valid	Trustworthy-invalid	Untrustworthy-valid	Untrustworthy-invalid
Experiment 1_original	RT	M	644.50	656.07	651.60	653.75
		SD	77.77	82.54	80.42	80.10
	ACC	M	0.97	0.96	0.97	0.96
		SD	0.04	0.02	0.03	0.03
Experiment 1_additional	RT	M	639.39	661.80	641.38	660.13
		SD	60.55	59.89	59.68	49.12
	ACC	M	0.98	0.98	0.98	0.98
		SD	0.02	0.01	0.02	0.02
Experiment 2_original	RT	M	623.89	632.74	633.41	620.99
		SD	63.41	66.31	63.67	64.88
	ACC	M	0.96	0.96	0.97	0.97
		SD	0.06	0.05	0.06	0.03
Experiment 2_additional	RT	M	635.18	653.44	644.06	651.52
		SD	45.40	51.68	47.49	51.02
	ACC	M	0.99	0.98	0.99	0.99
		SD	0.01	0.02	0.01	0.02

Descriptive data for original and newly recruited sample of Experiments 1 and Experiment 2 is presented, encompassing RT (ms) and ACC statistics

Table 3 ANOVA Results for ACC

	Main effect of validity	Main effect of trust	Interaction effect
Experiment 1_original	$F(1, 23) = 3.787, p = 0.064, \eta_p^2 = 0.141, BF_{01} = 1.449$	$F(1, 23) = 0.009, p = 0.923, \eta_p^2 < 0.001, BF_{01} = 4.049$	$F(1, 23) = 0.328, p = 0.572, \eta_p^2 = 0.014, BF_{01} = 2.571$
Experiment 1_additional	$F(1, 11) = 0.014, p = 0.908, \eta_p^2 = 0.001, BF_{01} = 2.427$	$F(1, 11) = 0.355, p = 0.563, \eta_p^2 = 0.031, BF_{01} = 1.880$	$F(1, 11) = 0.321, p = 0.587, \eta_p^2 = 0.028, BF_{01} = 2.364$
Experiment 2_original	$F(1, 23) = 0.059, p = 0.810, \eta_p^2 = 0.003, BF_{01} = 2.203$	$F(1, 23) = 3.796, p = 0.064, \eta_p^2 = 0.142, BF_{01} = 1.052$	$F(1, 23) = 2.676, p = 0.116, \eta_p^2 = 0.104, BF_{01} = 1.019$
Experiment 2_additional	$F(1, 11) = 0.774, p = 0.398, \eta_p^2 = 0.066, BF_{01} = 1.919$	$F(1, 11) = 2.210, p = 0.165, \eta_p^2 = 0.167, BF_{01} = 1.420$	$F(1, 11) = 0.171, p = 0.687, \eta_p^2 = 0.015, BF_{01} = 2.46$

Table 4 ANOVA Results for RT (ms)

	Main effect of validity	Main effect of trust	Interaction effect	GCE for trustworthy	GCE for untrustworthy
Experiment 1_original	valid vs. invalid: 650 vs. 659 ms, MSE=280.084, $F(1, 23) = 7.012, p = 0.014, \eta_p^2 = 0.234, BF_{10} = 3.302$	trustworthy vs. untrustworthy: 655 vs. 654 ms, MSE=175.770, $F(1, 23) = 0.027, p = 0.871, \eta_p^2 = 0.001, BF_{01} = 3.731$	MSE=136.113, $F(1, 23) = 5.625, p = 0.026, \eta_p^2 = 0.197, BF_{10} = 3.670$	15 ms, $t(23) = 3.529, p = 0.006, \text{Cohen's } d = -0.189$	3 ms, $t(23) = 0.816, p = 0.419, \text{Cohen's } d = 0.044$
Experiment 1_additional	valid vs. invalid: 640 vs. 661 ms, MSE=836.914, $F(1, 11) = 6.073, p = 0.031, \eta_p^2 = 0.356, BF_{10} = 2.76$	trustworthy vs. untrustworthy: 651 vs. 651 ms, MSE=119.161, $F(1, 11) = 0.002, p = 0.961, \eta_p^2 < 0.001, BF_{01} = 2.865$	MSE=214.943, $F(1, 11) = 0.187, p = 0.674, \eta_p^2 = 0.017, BF_{01} = 2.525$	22 ms, $t(11) = 2.394, p = 0.174, \text{Cohen's } d = 0.390$	19 ms, $t(11) = 2.003, p = 0.186, \text{Cohen's } d = 0.326$
Experiment 2_original	valid vs. invalid: 629 vs. 627 ms, MSE=166.041, $F(1, 23) = 0.463, p = 0.503, \eta_p^2 = 0.020, BF_{01} = 3.236$	trustworthy vs. untrustworthy: 628 vs. 627 ms, MSE=149.111, $F(1, 23) = 0.200, p = 0.659, \eta_p^2 = 0.009, BF_{01} = 3.571$	MSE=101.993, $F(1, 23) = 26.612, p < 0.001, \eta_p^2 = 0.536, BF_{10} = 1.568 \times 10^4$	9 ms, $t(23) = 2.657, p < 0.034, \text{Cohen's } d = 0.137$	-12 ms, $t(23) = -3.718, p = 0.003, \text{Cohen's } d = 0.192$
Experiment 2_additional	valid vs. invalid: 640 vs. 652 ms, MSE=213.732, $F(1, 11) = 9.281, p = 0.011, \eta_p^2 = 0.458, BF_{10} = 12.031$	trustworthy vs. untrustworthy: 644 vs. 648 ms, MSE=40.109, $F(1, 11) = 3.612, p = 0.084, \eta_p^2 = 0.247, BF_{10} = 1.095$	MSE=109.330, $F(1, 11) = 3.197, p = 0.101, \eta_p^2 = 0.225, BF_{01} = 1.029$	18 ms, $t(11) = 3.518, p = 0.013, \text{Cohen's } d = 0.373$	7 ms, $t(11) = 1.438, p = 0.332, \text{Cohen's } d = 0.152$

A: Results for original and additional sample in Experiment 1 and 2

B: Potential influence of sample status

In our methodology, we detailed our recruitment process, which involved two stages: initially recruiting 24 participants, followed by the recruitment of an additional 12 participants to enhance the robustness of our results. To mitigate

potential biases arising from sample recruitment timing, we incorporated sample status as a between-subject factor while maintaining trustworthiness and gaze-cue type as within-subject factors in both Experiment 1 and Experiment 2.

In Experiment 1, the three-way interaction did not reach statistical significance [MSE = 161.617, $F(1,34) = 0.722, p = 0.402, \eta_p^2 = 0.021, BF_{01} = 2.558$]. Likewise, in Experiment 2, the three-way interaction did not yield significance [MSE = 104.367, $F(1,34) = 2.103, p = 0.156, \eta_p^2 = 0.058, BF_{10} = 1.456$]. This suggests that irrespective of whether

participants were initially recruited or added subsequently, the interaction between trustworthiness and gaze-cue type exhibited a consistent pattern both in Experiment 1 and Experiment 2.

C: Joint analyses between experiment

Experiment 1 and Experiment 2

We conducted additional joint analyses by integrating data from both Experiment 1 and Experiment 2, treating Experiment itself as an additional between-subject factor, while retaining trustworthiness and gaze-cue type as within-subject factors. This approach enabled us to explore potential variations in the interaction between trustworthiness and gaze-cue type across the two experiments. However, the three-way interaction did not attain statistical significance [$MSE = 133.993$, $F(1, 70) = 2.736$, $p = 0.103$, $\eta_p^2 = 0.038$, $BF_{01} = 1.346$].

Experiment 3a and Experiment 3b

Similar to the joint analysis between Experiment 1 and 2, we aggregated data from Experiment 3a and 3b to investigate potential variations in the interaction across experiments. The three-way interaction was not found to be significant [$MSE = 149.212$, $F(1, 58) = 0.829$, $p = 0.366$, $\eta_p^2 = 0.014$, $BF_{01} = 2.915$], indicating a consistent pattern of null effect regarding the interaction between trustworthiness and gaze-cue type.

Acknowledgements This research was supported by grants from Yanliang Sun, including the Natural Science Foundation of Shandong Province (ZR2023MC204). Support from Xiaowei Ding included the National Natural Science Foundation of China (32271103), the Natural Science Foundation of Guangdong Province (2023A1515011762), the Guangzhou Science and Technology Plan Project - Leading Elite Program (2024A04J3301), and the Fundamental Research Funds for the Central Universities under the Major Project Cultivation and Emerging Interdisciplinary Cultivation Plan (24wkjc02). Additionally, Ding-guo Gao's support included the Natural Science Foundation of Guangdong Province (2024A1515010826) and the National Natural Science Foundation of China (32171073).

Author contributions Xiaowei Ding: Conceptualization; investigation; writing—review and editing; supervision. Jing Gan: Data collection; formal analysis; investigation; writing—original draft. Luzi Xu: Formal analysis; data collection. Xiaozhi Zhou: Methodology; Ding-guo Gao: Conceptualization; supervision. Yanliang Sun: Conceptualization; funding acquisition; project administration; resources; supervision.

Data availability The data will be available upon the acceptance of the paper for publication.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval Approval was obtained by the Research Ethics Board of Shandong Normal University (for experiment 1) and Research Ethics Board of Sun Yat-sen University (for experiment 2 and 3).

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge U Press.
- Ashraf, N., Bohnet, I., & Piankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics*, 9(3), 193–208. <https://doi.org/10.1007/s10683-006-9122-4>
- 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>
- Bayless, S. J., Glover, M., Taylor, M. J., & Itier, R. J. (2011). Is it in the eyes? Dissociating the role of emotion and perceptual features of emotionally expressive faces in modulating orienting to eye gaze. *Visual Cognition*, 19(4), 483–510. <https://doi.org/10.1080/13506285.2011.552895>
- Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review*, 13(6), 1061–1066. <https://doi.org/10.3758/BF03213926>
- 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>
- Ben-Ner, A., & Halldorsson, F. (2010). Trusting and trustworthiness: What are they, how to measure them, and what affects them. *Journal of Economic Psychology*, 31(1), 64–79. <https://doi.org/10.1016/j.joep.2009.10.001>
- Butler, J. K., & Cantrell, R. S. (1984). A behavioral decision theory approach to modeling dyadic trust in superiors and subordinates. *Psychological Reports*, 55(1), 19–28. <https://doi.org/10.2466/pr0.1984.55.1.19>
- Capozzi, F., Bayliss, A. P., & Ristic, J. (2018). Gaze following in multiagent contexts: Evidence for a quorum-like principle. *Psychonomic Bulletin & Review*, 25, 2260–2266. <https://doi.org/10.3758/s13423-018-1464-3>
- Capozzi, F., Bayliss, A. P., & Ristic, J. (2021). Standing out from the crowd: Both cue numerosity and social information affect attention in multi-agent contexts. *Quarterly Journal of Experimental Psychology*, 74(10), 1737–1746. <https://doi.org/10.1177/17470218211013028>
- Capozzi, F., Becchio, C., Willemsse, C., & Bayliss, A. P. (2016). Followers are not followed: Observed group interactions modulate subsequent social attention. *Journal of Experimental Psychology: General*, 145(5), 531–535. <https://doi.org/10.1037/xge0000167>
- Clifford, C. W. G., & Palmer, C. J. (2018). Adaptation to the direction of others' gaze: A review. *Frontiers in Psychology*, 9, 2165. <https://doi.org/10.3389/fpsyg.2018.02165>
- Cole, G. G., Smith, D. T., & Atkinson, M. A. (2015). Mental state attribution and the gaze cueing effect. *Attention, Perception, & Psychophysics*, 77(4), 1105–1115. <https://doi.org/10.3758/s13414-014-0780-6>
- Couch, L. L., & Jones, W. H. (1997). Measuring Levels of Trust. *Journal of Research in Personality*, 31(3), 319–336. <https://doi.org/10.1006/jrpe.1997.2186>
- Dalmaso, M., Castelli, L., & Galfano, G. (2020). Social modulators of gaze-mediated orienting of attention: A review. *Psychonomic*

- Bulletin & Review*, 27(5), 833–855. <https://doi.org/10.3758/s13423-020-01730-x>
- Dalmaso, M., Pavan, G., Castelli, L., & Galfano, G. (2011). Social status gates social attention in humans. *Biology Letters*, 8(3), 450–452. <https://doi.org/10.1098/rsbl.2011.0881>
- Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science*, 3(5), 562–571. <https://doi.org/10.1177/1948550611430272>
- Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers visuospatial orienting by adults in a reflexive manner. *Visual Cognition*, 6(5), 509–540. <https://doi.org/10.1080/135062899394920>
- Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, 24(6), 581–604. [https://doi.org/10.1016/S0149-7634\(00\)00025-7](https://doi.org/10.1016/S0149-7634(00)00025-7)
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77–83. <https://doi.org/10.1016/j.tics.2006.11.005>
- Freitag, M., & Traunmüller, R. (2009). Spheres of trust: An empirical analysis of the foundations of particularised and generalised trust. *European Journal of Political Research*, 48(6), 782–803. <https://doi.org/10.1111/j.1475-6765.2009.00849.x>
- 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>
- Gawronski, B., & Quinn, K. A. (2013). Guilty by mere similarity: Assimilative effects of facial resemblance on automatic evaluation. *Journal of Experimental Social Psychology*, 49(1), 120–125. <https://doi.org/10.1016/j.jesp.2012.07.016>
- Gibson, J. J. (1979). The ecological approach to visual. *Perception*, 14, 332. <https://doi.org/10.4324/9781315740218>
- Gray, K. L. H., Haffey, A., Mihaylova, H. L., & Chakrabarti, B. (2018). Lack of privileged access to awareness for rewarding social scenes in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 48(10), 3311–3318. <https://doi.org/10.1007/s10803-018-3595-9>
- Hale, J., Payne, M. E., Taylor, K. M., Paoletti, D., & De Hamilton, A. F. (2018). The virtual maze: A behavioural tool for measuring trust. *Quarterly Journal of Experimental Psychology*, 71(4), 989–1008. <https://doi.org/10.1080/17470218.2017.1307865>
- Hermens, F., & Walker, R. (2012). Do you look where I look? Attention shifts and response preparation following dynamic social cues. *Journal of Eye Movement Research*. <https://doi.org/10.16910/jemr.5.5.5>
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 65–70.
- JASP Team (2024). JASP (Version 0.18.3)[Computer software].
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.
- Jessen, S., & Grossmann, T. (2020). Neural evidence for the impact of facial trustworthiness on object processing in a gaze-cueing task in 7-month-old infants. *Social Neuroscience*, 15(1), 74–82. <https://doi.org/10.1080/17470919.2019.1651764>
- Jones, B. C., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., Arinze, N. C., Ndakaihe, I. L. G., Bloxson, N. G., Lewis, S. C., Foroni, F., Willis, M. L., Cubillas, C. P., Vadillo, M. A., Turiegano, E., Gilead, M., Simchon, A., Saribay, S. A., Owsley, N. C., Jang, C., & Coles, N. A. (2021). To which world regions does the valence-dominance model of social perception apply? *Nature Human Behaviour*, 5(1), 159–169. <https://doi.org/10.1038/s41562-020-01007-2>
- Jones, B. C., DeBruine, L. M., Main, J. C., Little, A. C., Welling, L. L. M., Feinberg, D. R., & Tiddeman, B. R. (2010). Facial cues of dominance modulate the short-term gaze-cueing effect in human observers. *Proceedings of the Royal Society B: Biological Sciences*, 277(1681), 617–624. <https://doi.org/10.1098/rspb.2009.1575>
- 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(4), 476–485. <https://doi.org/10.1521/soco.2011.29.4.476>
- Kingstone, A., Kachkovski, G., Vasilyev, D., Kuk, M., & Welsh, T. N. (2019). Mental attribution is not sufficient or necessary to trigger attentional orienting to gaze. *Cognition*, 189, 35–40. <https://doi.org/10.1016/j.cognition.2019.03.010>
- Klapper, A., Dotsch, R., van Rooij, I., & Wigboldus, D. H. J. (2016). Do we spontaneously form stable trustworthiness impressions from facial appearance? *Journal of Personality and Social Psychology*, 111(5), 655–664. <https://doi.org/10.1037/pspa0000062>
- Kleinke, C. L. (1986). Gaze and eye contact: a research review. *Psychological Bulletin*, 100(1), 78–100. <https://doi.org/10.1037/0033-2909.100.1.78>
- Kompatsiari, K., Ciardo, F., & Wykowska, A. (2022). To follow or not to follow your gaze: The interplay between strategic control and the eye contact effect on gaze-induced attention orienting. *Journal of Experimental Psychology. General*, 151(1), 121–136. <https://doi.org/10.1037/xge0001074>
- Kuhn, G., & Tipples, J. (2011). Increased gaze following for fearful faces. It depends on what you're looking for. *Psychonomic Bulletin & Review*, 18(1), 89–95. <https://doi.org/10.3758/s13423-010-0033-1>
- Langton, S. R. (2009). Gaze perception: Is seeing influenced by believing? *Current Biology*, 19(15), R651–R653. <https://doi.org/10.1016/j.cub.2009.06.016>
- Langton, S. R. H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, 6(5), 541–567. <https://doi.org/10.1080/135062899394939>
- Langton, S. R., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, 4(2), 50–59.
- Law, A. S., Langton, S. R., & Logie, R. H. (2010). Assessing the impact of verbal and visuospatial working memory load on eye-gaze cueing. *Visual Cognition*, 18(10), 1420–1438. <https://doi.org/10.1080/13506285.2010.496579>
- Ma, F., Xu, F., & Luo, X. (2016). Children's facial trustworthiness judgments: Agreement and relationship with facial attractiveness. *Frontiers in Psychology*, 7, Article 499. <https://doi.org/10.3389/fpsyg.2016.00499>
- Marzi, T., Righi, S., Ottonello, S., Cincotta, M., & Viggiano, M. P. (2014). Trust at first sight: Evidence from ERPs. *Social Cognitive and Affective Neuroscience*, 9(1), 63–72. <https://doi.org/10.1093/scan/nss102>
- McKay, K. T., Grainger, S. A., Coundouris, S. P., Skorich, D. P., Phillips, L. H., & Henry, J. D. (2021). Visual attentional orienting by eye gaze: A meta-analytic review of the gaze-cueing effect. *Psychological Bulletin*, 147(12), 1269–1289. <https://doi.org/10.1037/bul0000353>
- McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *The Academy of Management Review*, 23(3), 473–490. <https://doi.org/10.2307/259290>
- 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>

- Myllyneva, A., & Hietanen, J. K. (2015). There is more to eye contact than meets the eye. *Cognition*, *134*, 100–109. <https://doi.org/10.1016/j.cognition.2014.09.011>
- 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*, 11087–11092. <https://doi.org/10.1073/pnas.0805664105>
- Petrican, R., English, T., Gross, J. J., Grady, C., Hai, T., & Moscovitch, M. (2012). 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>
- Rule, N. O., Ambady, N., & Hallett, K. C. (2009). Female sexual orientation is perceived accurately, rapidly, and automatically from the face and its features. *Journal of Experimental Social Psychology*, *45*(6), 1245–1251. <https://doi.org/10.1016/j.jesp.2009.07.010>
- Sato, W., Okada, T., & Toichi, M. (2007). Attentional shift by gaze is triggered without awareness. *Experimental Brain Research*, *183*(1), 87–94. <https://doi.org/10.1007/s00221-007-1025-x>
- Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*(9), 668–671. <https://doi.org/10.1016/j.cub.2008.03.059>
- Shi, J., Weng, X., He, S., & Jiang, Y. (2010). Biological motion cues trigger reflexive attentional orienting. *Cognition*, *117*(3), 348–354. <https://doi.org/10.1016/j.cognition.2010.09.001>
- Singular Inversions (2006). FaceGen 3.1 Full SDK Documentation. <http://facegen.com>. Accessed 5 Jun 2007.
- 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*(7–8), 802–814. <https://doi.org/10.1080/13506285.2017.1338321>
- 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>
- Teufel, C., Alexis, D. M., Clayton, N. S., & Davis, G. (2010a). Mental-state attribution drives rapid, reflexive gaze following. *Attention, Perception, & Psychophysics*, *72*(3), 695–705. <https://doi.org/10.3758/APP.72.3.695>
- Teufel, C., Fletcher, P. C., & Davis, G. (2010b). Seeing other minds: Attributed mental states influence perception. *Trends in Cognitive Sciences*, *14*(8), 376–382. <https://doi.org/10.1016/j.tics.2010.05.005>
- Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008a). Evaluating face trustworthiness: A model based approach. *Social Cognitive & Affective Neuroscience*, *3*(2), 119–127. <https://doi.org/10.1093/scan/nsn009>
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623–1626. <https://doi.org/10.1126/science.1110589>
- Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*(1), 519–545. <https://doi.org/10.1146/annurev-psych-113011-143831>
- Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008b). 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. (1995). Joint attention as social cognition. *Joint Attention: Its Origins and Role in Development*, *103130*, 103–130.
- Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. <https://doi.org/10.3758/s13423-017-1343-3>
- Westfall, J. (2015). PANGEA: Power analysis for general ANOVA designs. 4. Unpublished manuscript. Retrieved from <http://jakewestfall.org/publications/pangea.pdf>.
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, *37*(3), 395–412. <https://doi.org/10.1037/0022-3514.37.3.395>
- 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>
- Zohary, E., Harari, D., Ullman, S., Ben-Zion, I., Doron, R., Attias, S., Porat, Y., Sklar, A. Y., & Mckyton, A. (2022). Gaze following requires early visual experience. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(20), e2117184119. <https://doi.org/10.1073/pnas.2117184119>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.