Gaze towards my choice: Noneconomic social interaction changes interpersonal trust only with positive feedback

Ziyi Duan¹ · Tian Ye² · Anita Poggi³ · Xiaowei Ding¹

Published online: 4 August 2020
© The Psychonomic Society, Inc. 2020

Abstract

Trust is a crucial foundation of human society, dynamically changing along with interpersonal interactions. Previous research suggested that both gaze behaviors and affective states might influence our trust evaluation. Here, we created an uncertain decision-making environment to investigate how these two aspects change interpersonal trust during social interactions. Specifically, participants made a forced choice between two price-matched goods to determine which one was more expensive. After participants made a choice, the face in the center of the screen would gaze towards or away from the participants' choice, and then participants would receive timely feedback informing whether their choice was correct or not. In Experiment 1 ($N = 34$), we found people increased their trust ratings for gaze-followers but reduced them for gaze-unfollowers when they received positive feedback, but stayed with their initial judgments when receiving negative feedback. Experiment 2 ($N = 20$) and Experiment S1 ($N = 20$) further revealed the same change patterns on participants' trust investments in different conditions. These results suggest that gaze-following behavior can change people's trust evaluations as well as trust behaviors, while affective states induced by different feedback modulate this process in a flexible manner.

Keywords Social interaction · Trust · Gaze-following · Decision-making

Trust is a nebulous construct and a core component of social life, especially for promoting cooperation and strengthening ingroup bonds (Krueger & Meyer-Lindenberg, 2019; Lount, 2010; Robbins, 2016). Although we seem to be experts on quickly evaluating others' trustworthiness (Willis & Todorov, 2006), such judgment is not always reliable as people are also good at deceiving and camouflaging (Rule, Krendl, Ivcevic, & Ambady, 2013). Such complexity of social interaction puts forward the challenge of fast

updating our trustworthiness evaluations. Nonetheless, we know little about how trust evaluations change during interpersonal interactions (Tuk, Verlegh, Smidts, & Wigboldus, 2009), within which most research focuses on economic behaviors while overlooking noneconomic social interactions.

Social interactions in economic life are composed of gains and losses, hence monetary factors are key elements in changing trustworthiness evaluations (Rule et al., 2013). However, economic gains and losses are not the sole origin of trust. With people we meet for the first time—for example, random tourists from a day trip—our trust in them is likely to be influenced by noneconomic interactions during initial contact. Do they agree to go with us at a crossroad? Do we in the end make a correct decision? Although previous research on trust rarely linked trust formulation and change with these routine social interactions, recent studies suggested that social signals people perceived in such scenarios could change our person perception, which might include trust evaluations (Singer, Kiebel, Winston, Dolan, & Frith, 2004; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). However, it is largely unknown whether, and how, such social signals can alter our trust judgments.

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13423-020-01785-w>) contains supplementary material, which is available to authorized users.

✉ Xiaowei Ding
dingxw3@mail.sysu.edu.cn

- ¹ Guangdong Provincial Key Laboratory of Social Cognitive Neuroscience and Mental Health, and Department of Psychology, Sun Yat-Sen University, Guangzhou 510006, People's Republic of China
- ² Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AZ, UK
- ³ Department of Psychology, University of Milano-Bicocca, Milan, Italy

Because of its capacity to convey a wealth of information about people's mental states, gaze is considered as a rudimentary form of social signals (Frischen, Bayliss, & Tipper, 2007). Recent studies showed that gaze could directly change our perception of other persons, like facial preference (Kuzmanovic et al., 2009) and trustworthiness (Bayliss & Tipper, 2006). However, in these and many similar studies, participants were passive observers of other's gaze, rather than active gaze initiators who use gaze to lead other's attention (Bayliss et al., 2013). Consider the day trip example. When we are the decision-makers in such cases, how do we develop trust in our partner when we take the lead of interactions?

In the current study, we are interested in asking whether another form of gaze, other's gaze-following behavior, can change our trust ratings of them. Gaze-following refers to the process in which others shift their gaze towards the target we initially attend to (Driver et al., 1999; Friesen & Kingstone, 1998). Previous studies have hinted that people seem to have internal templates in which other's gaze-following behavior affects our impression of them. Individuals who constantly unfollow our gaze would be regarded as norm violators (e.g., Bayliss et al., 2013; Edwards, Stephenson, Dalmaso, & Bayliss, 2015). Bayliss and colleagues asked participants to select a preferred object from two alternatives, and once the participant's preference was detected by the computer, a face on the screen would gaze towards or away from it. They found that participants thought gaze-followers were more likable than gaze-unfollowers, indicating the impact of gaze-following behavior on our affective evaluations. Although affective evaluation shares a close link with trust evaluation (Frischen et al., 2007), it has not yet been investigated whether such gaze-following behavior can also influence interpersonal trust.

Despite the lack of direct examination on the relationship between gaze-following behavior and trust evaluation, we can predict the results based on previous research. For example, in the classical gaze-cuing paradigm, faces providing valid gaze cue were perceived as being more trustworthy (Mason, Tatkov, & Macrae, 2005; Strachan, Kirkham, Manssuer, Over, & Tipper, 2017; Strachan, & Tipper, 2017). Moreover, according to Bayliss and colleagues' work (2013; Bayliss & Tipper, 2006), gaze-following behavior increased preference to the face, while gaze-unfollowers were less preferred. Therefore, it is natural to hypothesize that gaze-followers will receive more trust while gaze-unfollowers may lose trust after social interactions. Nevertheless, recent evidence supports a strong and complex relationship between trust and affect. For instance, Hommel and Colzato (2015) proposed that interpersonal trust is modulated by the truster's current affective states. Moreover,

even in the original study on the gaze-trust effect (Bayliss, Griffiths, & Tipper, 2009), the effect of gaze on interpersonal trust existed only when the faces created a positive social context by showing a smile, but not when the faces showed negative or neutral expressions. To further show the relationship between gaze and trust, we thus imported another factor that could modulate participants' moods: the feedback of a decision made with uncertainty. In cases where people have made a difficult decision, the feedback of the decision can cause a temporary change in the decision-maker's current mood (Steenbergen, Band, & Hommel, 2009). Previous studies showed that this process might be related to the impact of such feedback on dopamine level (Schott et al., 2008; Schultz, 2007). Therefore, providing positive feedback could temporarily induce participants' positive feelings while negative feedback could cause negative moods.

Researchers also proposed theoretical models to explain the dissociable affective effect on trust evaluations (Fiedler, 2001; Fiedler, Nickel, Asbeck, & Pagel, 2003). According to the accommodation-assimilation model, people in positive mood tend to rely on heuristic process and use activated schemas (e.g., stereotypes) when evaluating others, yet when in negative mood they switch to systematic thinking and become more reflective during evaluation. In terms of trust ratings, we thus hypothesize that when participants are in positive mood, they may rely on the heuristic cues available in this context to make judgments—whether the trustee followed our choice or not. So, the heuristic thinking process drives participants to increase their trust in gaze-followers while decreasing trust in gaze-unfollowers. In contrast, when negative feelings are invoked by negative feedback, people may turn to deliberate thinking and give more conservative judgement about others' trustworthiness.

To test these hypotheses, in the current study, we created a social interactive scene by combining the gaze-following paradigm (Bayliss et al., 2013) with an uncertain choice task. In this task, participants were required to choose the more expensive object between two price-matched alternatives. Once participants made a choice, a face on the screen would gaze towards or away from participant's choice. Then participants would receive a correct or incorrect feedback about their choice at the end of each trial. The two types of gaze-following behaviors were combined with two types of feedback, resulting in four different groups of interactive partners, on which participants need to make trust evaluations. To further test whether such trust-rating changes can extend to real trust-behavior changes, we added an economic game (the Trust Game) in Experiment 2 and S1 (see Supplementary), and predicted a similar result pattern on participants' economic behaviors.

Experiment 1

Method

Participants

Thirty-four participants (19 females, mean age = 20.2 years, $SD = 2.1$ years) from Sun Yat-Sen University were recruited. All participants had normal or corrected-to-normal visual acuity, and no reported syndrome of face blindness. Postexperiment surveys indicated that they were naïve to the experimental purpose. Considering the severe replication crisis in psychology (e.g., Pashler & Wagenmakers, 2012), we used the default effect size suggested by PANGEA (generalized Cohen's $d = 0.45$, which equals to $\eta_p^2 = 0.05$; Westfall, 2015) for the Gaze-Following Behavior $\times$ Feedback interaction in a within-subjects design. With 95% power at a .05 significance level, PANGEA suggested a sample size of 34 participants. All experimental procedures in our current study were approved by the Sun Yat-Sen University Institutional Review Board.

Stimulus selection

Because we need to create a decision environment filled with uncertainty, and face attractiveness might confound with how much we trust others (R. K. Wilson & Eckel, 2006), before we started the formal experiment, we asked a different group of participants to estimate the price for each paired objects and the attractiveness for each face. Twelve undergraduate students (six females, mean age = 20.5 years, $SD = 1.5$) were recruited. Participants all had normal or corrected-to-normal visual acuity and, importantly, no facial blindness. Participants received 10RMB for their participation.

Face selection

Participants were asked to rate the attractiveness of 26 smiling faces, which were selected from the standard face library NimStim Set of Facial Expressions (Tottenham et al., 2009). For each trial, participants were asked to rate the attractiveness of one face by adjusting a rating bar below each face. The rating range was from 1 to 15, where 1 means the lowest attractiveness and 15 means the highest attractiveness. As a result, we selected eight faces with medium attractiveness (mean = 6.72, $SD = 2.12$), four males and four females, which were then randomly divided into four groups (see Fig. 1 and Table 1). To ensure the faces' attractiveness was balanced among different groups, we computed each face group's average attractive ratings (two faces in a group) and compared that of the four groups by using a repeated-measures analysis of variance (ANOVA) with Greenhouse–Geisser correction. Results demonstrated that the faces' attractiveness did not

show significant differences among different groups, $F(1.63, 17.92) = 0.52$, $p = .57$, $\eta_p^2 = 0.045$.

After selecting faces, three versions of each face were produced, one with direct gaze (i.e., the person in the picture will look directly towards the participant when the picture is presented at the middle of the screen), one with the person looking to the left, and another looking to the right. Participants seated at a distance of approximately 60 cm in front of the screen, and the faces measured $5.0^\circ \times 6.5^\circ$ (visual angle). Eye regions measured between 2.0° and 2.5° from the left corner of the left eye and the right corner of the right eye. The eyes measured $0.2^\circ \times 0.5^\circ$, with pupils/irises of $0.2^\circ \times 0.2^\circ$. According with Bayliss and colleagues (2013), all faces showed a moderate smiling expression and were initially presented looking straight ahead. The manipulations of the faces consisted in making the eyes to appear to look towards the left or look towards the right.

Object selection

One hundred pairs of pictures with price-matched objects were chosen from Brady, Konkle, Alvarez, and Oliva's (2008) study. For each trial, two paired objects were presented on either side of the screen, and participants were then asked to select which one was more expensive. Each picture measured $5.2^\circ \times 5.2^\circ$, located 5.2° away from the central face. We recorded participant's choice from each trial, then calculated the selection rates of objects in each pair by each person. Four pairs provided biased results, which were removed from our materials for the formal experiment, as all participants selected the same object in these pairs. We divided the remained 96 paired objects into two groups and calculated the selection rate of each object. We then we compared the selection rates of objects in each pair by using a paired-samples t test. Results indicated that participants had no significant preference in those pairs, $t(95) = -1.270$, $p = .207$, Cohen's $d = -0.029$.

Procedure

Participants arrived at the lab and were presented consent forms to sign. Then they were guided to sit naturally in front of a 23.8-inch FHD monitor with a distance of approximately 60 cm. Task stimuli were displayed on a white (RGB: 255, 255, 255) background. The refresh rate was 60 Hz, and the screen resolution was $1,920 \times 1,080$ pixels. We used MATLAB (The MathWorks, Inc., Natick, MA) with Psychophysics Toolbox extensions (Brainard, 1997) to program and run the experiment. The whole procedure was divided into three sections as follows (pretrust rating, price-comparing task, and posttrust rating), and participants were instructed to complete each section sequentially (see Fig. 2).

Fig. 1 A pictorial description of the 2×2 conditions. In this example, the participant selected the purple guitar on the left, and the face in the center would gaze towards (follower) or away (unfollower) from the purple guitar (the dashed square and arrows were only for demonstration and

did not exist in the experiment). At the end of each trial, participants would receive feedback stating their choices were either correct or incorrect. (Color figure online)

Pre-trust rating

To capture how participants changed their trust ratings after different social interactions, we adopted the pretrust and posttrust rating methods. Participants were first told that they need to rate the trustworthiness for different faces. Eight faces were presented in serial with random order, and participants were asked to rate how trustworthy they found each face on a 1 (*highly untrustworthy*) to 50 (*highly trustworthy*) scale. Participants rated the faces' trustworthiness by adjusting the rating bar below each face, which was initially set in the middle 25. There was no time limit for the prerating, and after participants confirmed their ratings by pressing the “space” key, the next face would immediately appear.

Price-comparing task

After the pretrust rating, participants were told that they needed to do a price-comparing task. At the start of each

trial, participants fixated a central cross for ~500–1,000 ms. Then, the cross was replaced by a face, and two paired objects were presented on either side of it. Participants were asked to guess which object was more expensive, without a time limit. They were instructed to press the “F” key to choose the object on the left, and press the “J” key to choose the object on the right. Once they made a choice, after a 500-ms delay, the face in the center of the screen would look to the left or right. Participants were required to remember this choice, and press the “space” key to confirm. After that, participants received “correct” or “incorrect” feedback for 2,000 ms, based on the trial condition. To test whether participants paid enough attention to the person's choice or not, we inserted 30 catch trials into the formal trials. On each of the catch trial, we asked participants to choose the object that the person on the screen had selected, then gave them immediate feedback based on their answer. The accuracy of the memory task was recorded to exclude participants who did not

Table 1 A detailed description of the 2×2 conditions in all experiments

Group	Detailed description
Correct & follower	Faces always gazed towards participant's choices, and participants always received correct feedback.
Correct & unfollower	Faces always gazed away from participant's choices, and participants always received correct feedback.
Incorrect & follower	Faces always gazed towards participant's choices, and participants always received incorrect feedback.
Incorrect & unfollower	Faces always gazed away from participant's choices, and participants always received incorrect feedback.

Fig. 2 Procedure framework of Experiment 1, including three sections. The first section was pre-trust rating. The second section was the price-comparing task. The last section was post-trust rating. (Color figure online)

attend enough to the task, but this was not given to the participants. To further ensure that participants focused on the task, we also informed them before the experiment that their accuracy on choosing the more expensive object decided the final payment they would receive. If the final correct rate was equal to 50%, participants would receive 15 RMB (approximately US\$2.18) as the payment; if less than 50%, they would receive 10 RMB (approximately US\$1.45); if higher than 50%, they would receive 20 RMB (approximately US\$2.91; see Fig. 2). However, in our designed task, all participants' real accuracy was set at 50%, and thus we randomly gave the final task accuracy feedback as between 45.8% and 54.2% at the end of the price-comparing task. At the end of the experiment, we explained this to all the participants.

We manipulated the gaze-following behavior of the person in the center of the screen and the feedback from each trial. Eight faces were then selected based on a previous screening procedure (see [Face selection](#) section, above) and were randomly assigned to 2×2 experimental conditions, with two faces appearing in each condition. The detailed description of the four conditions are listed in Table 1 (also see Fig. 1 for an illustration). This way, we guaranteed the behavioral consistency for each person, as one face would only appear in one condition. Faces in different conditions were balanced across all participants.

Participants completed 96 formal trials (three blocks of 32 trials), with another 30 catch trials randomly inserted. The formal trials consisted of 48 trials with "correct" feedback and 48 trials with "incorrect" feedback. Within each feedback condition, the faces in the center of the screen gazed towards participant's choices on half of all trials and away from for the other half. Each condition included two faces (one male and one female), so each face would appear 12 times (four times

for each block) throughout the whole experiment. Participants needed to complete four practice trials with two novel faces (one male and one female) before starting the formal experiment.

Post-trust rating

After the price-comparing task, participants were asked to re-do the trust rating on the previous eight faces. Stimuli and procedures were the same as in the pretrust rating section.

Data analysis

To test whether different interactive modes can influence people's trust evaluations of others, we calculated differential trust scores by subtracting participants' prerating scores from the postrating scores for the same face. As each interactive condition used two faces, we first calculated the differential scores individually, then averaged them and used the average score as the index of trust change in the four different conditions.

To test whether any significant trust evaluation changed before and after interactions, we first performed one-sample *t* tests (Bonferroni corrected) on the differential trust scores for all the four conditions, to reveal whether any differential trust score was significantly different from zero. To compare whether any significant difference on trust changed across the four conditions, we used the two-way repeated-measures ANOVA, with gaze-following and feedback as within-subjects factors. All statistical analyses were conducted with R 4.0.1 (R Core Team, 2017) and SPSS Statistics 22 (IBM, NY, USA).

Results

Average accuracy rate on catch trials was 97.4% ($SD = 0.04$). This implies that most of the time our participants focused on the task, so all the data were included for further analyses.

Firstly, one-sample t tests showed that only with correct feedback did other's gaze-following or unfollowing behaviors induce significant changes on participants' trust ratings: for correct feedback & gaze-follower, $t(33) = 3.73$, $p = .001$, Cohen's $d = 0.64$, 95% CI [1.76, 6.00], and correct feedback & gaze-unfollower, $t(33) = -3.55$, $p = .001$, Cohen's $d = -0.61$, 95% CI [-4.77, -1.29]; for incorrect feedback & gaze-follower, $t(33) = -0.50$, $p = .62$, Cohen's $d = -0.09$, 95% CI [-4.02, 2.43], and incorrect feedback & gaze-unfollower, $t(33) = 1.12$, $p = 0.27$, Cohen's $d = 0.19$, 95% CI [-1.58, 5.43] (see Table 2). The two-way repeated-measures ANOVA on differential trust scores revealed a significant interactive effect between gaze-following and feedback, $F(1, 33) = 8.14$, $p = .007$, $\eta_p^2 = 0.20$, with no significant main effect, for feedback, $F(1, 33) = 0.02$, $p = .91$, $\eta_p^2 < 0.001$, and for gaze-following, $F(1, 33) = 2.59$, $p = .12$, $\eta_p^2 = 0.07$. Post-hoc test revealed that when participants received correct feedback, they would trust gaze-followers more compared with unfollowers, $t(33) = 5.06$, $p < .001$, Cohen's $d = 0.91$, 95% CI [4.13, 9.69]. Nonetheless, when people received negative feedback, their trust to followers and unfollowers showed no significant difference, $t(33) = -1.01$, $p = .32$, Cohen's $d = -0.36$, 95% CI [-8.19, 2.75] (see Fig. 3).

Discussion

In Experiment 1, we found that the influence from other's gaze-following behavior on our trust evaluations was modulated by task feedback. Compared with participants' initial trust rating, when participants received correct feedback, others' gaze-following behavior raised participants' trust ratings while unfollowing behavior decreased them. However, this opposite trend disappeared when participants received negative feedback. These results revealed a dissociable influence of gaze-following behavior on interpersonal trust, which confirmed our predictions when simultaneously considering the effect of gaze-following behavior and affective state.

Table 2 Descriptive statistics of differential trust scores and one-sample t test results in Experiment 1

Group	M	SD	t	p	95% CI
Correct & follower	3.88	6.07	3.73	.001	[1.76, 6.00]
Correct & unfollower	-3.03	4.98	-3.55	.001	[-4.77, -1.29]
Incorrect & follower	-0.79	9.24	-0.50	.62	[-4.02, 2.43]
Incorrect & unfollower	1.93	10.04	1.12	.27	[-1.58, 5.43]

Fig. 3 Differential trust scores in Experiment 1. All error bars in this paper represent 95% CI within subjects (Cousineau, 2005). Asterisks indicate significant differences compared with zero (one-sample t test) and between groups (repeated-measures ANOVA). ** $p < .01$. *** $p < .001$

Experiment 2

In Experiment 2, we added a one-shot Trust Game (Berg, Dickhaut, & McCabe, 1995) to further determine whether any change of trust ratings in Experiment 1 could affect participants' trust behaviors. In addition, as during the debriefing in Experiment 1 some participants reported that they could not remember their initial trust scores for some faces. It seems that interactions during the price-comparing task not only changed participants' postexperiment ratings on trust but also changed their impression of strangers' initial trustworthiness, which is likely to be a memory bias caused by different types of interactions. To test this possible subliminal phenomenon, we also asked participants to recall their initial trust score before giving their postratings.

Method

Participants

Twenty students (13 females, mean age = 21.8 years, $SD = 3.5$ years) from Sun Yat-Sen University were recruited. To obtain the comparable effect size of the two-way interaction (Gaze-Following Behavior $\times$ Feedback) as in Experiment 1 ($\eta_p^2 = 0.20$), we used PANGAEA to estimate the sample size, which suggested nine participants with 95% power at a .05 significance level. To ensure the reliability of our experiment, we raised the sample size to 20, which could provide a nearly 100% power at a .05 significance level. Moreover, in a supplementary experiment (see Supplementary), we used the same sample size ($N = 20$) to replicate our findings. All participants had normal or corrected-to-normal visual acuity, and

no one reported a history of face blindness. None of them took part in Experiment 1, and the postexperiment survey indicated that they were naïve to the experimental purpose.

Procedure

Participants arrived at the lab and were presented with the consent forms to sign. They then completed the pretrust rating and price-comparing task as in Experiment 1. Before entering the postrating stage, they were prompted to recall their initial trust rating scores on the face, and then gave a postrating score. They were then instructed to complete a single-round Trust Game (see Fig. 4 for the procedure).

Participants read an introduction of the rules of the Trust Game on the screen without time limit. In the on-screen instruction, they were told to imagine that they were playing a game with the person on the screen for a single round, and, as the initiator, they needed to decide how much to share between \$0 and \$50 with this person. The amount of money selected was multiplied by 4 and transferred to the other person, then he or she would decide how much to share back. In fact, the partners' responses were all preprogrammed (van't Wout & Sanfey, 2008). In half of the trials, the other person would send more money back, so in the end participants would have more money than before, but in the other half, the other person would share less or nothing, so participants would have less than before. Importantly, it was emphasized to the participants that the partner would play the game independently of each other and only once with the participants. Moreover, the order of the faces and the response pattern were all randomized across all participants, to minimize the possible interference of the partner's responses to the following trust evaluation. As there were eight different faces in the

price-comparing task, participants would play one round of the Trust Game with each of them, resulting in eight times total. The investment to each person was recorded as an indirect measurement of participants' trust of them.

Data analysis

Memory bias To test whether social interaction can lead to subliminal memory bias of initial trust scores, we calculated the deviation trust scores by subtracting the prerating scores from the recalled prerating scores and averaged them between the two faces within each condition. We first ran one-sample *t* tests (with Bonferroni correction) to test whether any deviation trust score was significantly different from zero. Then, a two-way repeated-measures ANOVA was performed to test different patterns among 2×2 conditions. To avoid the potential interference of the memory bias on our results, we used the recalled prerating scores as the baseline in the following data analyses.

Trust ratings We calculated the differential trust scores by subtracting the recalled prerating scores from the postrating scores for the same face, then averaged the two faces' differential scores within each condition. We performed the same one-sample *t* tests and two-way ANOVA on the differential trust scores.

Besides, to rule out the possible influence of the memory bias, we did extra analyses by using the initial prerating scores as baseline to compare trust rating changes, as Experiment 1. We also added the baseline (prerating vs. recalled prerating scores) as a new factor into the repeated-measures ANOVA to discover whether any effect existed (see [Supplementary](#) for more details).

Fig. 4 Procedure framework of Experiment 2, including three sections. The first two sections were exactly the same as Experiment 1, while the third section had two additional tasks. The recalling task was aimed to

exclude the memory bias, and the Trust Game was used to measure participants' trust behaviors. (Color figure online)

Table 3 Descriptive statistics of different scores in Experiment 2

Group	Deviation trust scores (memory bias)		Differential trust scores (trust change)		Differential invest scores		Standardized differential invest scores	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Correct & follower	1.15	3.88	5.45	9.38	1.95	11.97	0.44	1.21
Correct & unfollower	0.73	3.40	−5.18	9.71	−10.88	11.63	−0.39	1.26
Incorrect & follower	−0.78	3.51	−1.23	7.16	−7.85	11.54	−0.23	1.21
Incorrect & unfollower	−0.75	2.85	1.03	8.73	−4.65	12.26	0.18	1.14

Trust behaviors We aimed to explore whether the trust changes would influence the subsequent trust decisions in an economic game. Since the money invested by the investor could represent the investors' trust to the counterpart (Berg et al., 1995), we calculated the differential invest scores by subtracting the recalled prerating scores from the invested money and averaged them across two faces within each condition. Then, we conducted a two-way repeated-measures ANOVA to compare them across the 2×2 conditions.

Although both the invested money and prerating trust scores were used to measure participants' trust to interactive partners, some may argue that they might have different scales. Thus, to make the data more comparable and easily explained, we transformed the original data sets to standard *Z* scores and reran the analyses to ensure that our results were consolidated. Specifically, both the recalled prerating scores and participants' investments were converted into the standard *Z* scores, then the standardized differential invest scores were calculated (standardized investments − standardized recalled prerating scores). We then performed a two-way repeated-measures ANOVA on the standardized differential invest scores.

Results

Average accuracy rate on catch trials was 97.33% (*SD* = 0.03). This implies that most of the time our participants focused on the task, so all the data were included for further analyses, as in Experiment 1. All the descriptive statistics are listed in Table 3.

As for the memory bias, one-sample *t* tests showed null effects for all four conditions,¹ which indicated that participants' subjectively reported memory decay did not significantly influence their first trust impression. However, the

two-way ANOVA on deviation trust scores revealed a significant main effect of feedback, $F(1, 19) = 6.10, p = .02, \eta_p^2 = 0.24$, and no other main effect or interaction reached significance level: main effect for gaze-following, $F(1, 19) = 0.06, p = .81, \eta_p^2 = 0.003$; interactive effect, $F(1, 19) = 0.10, p = .76, \eta_p^2 = 0.005$. In other words, when receiving incorrect feedback, people tended to report lower initial trust ratings on their interactive partners compared with receiving correct feedback.

As for the differential trust scores, one-sample *t* tests indicated that only in the correct-feedback & gaze-follower condition, $t(19) = 2.60, p = .018$, Cohen's $d = 0.58$, 95% CI [1.06, 9.84], and correct-feedback & gaze-unfollower condition, $t(19) = -2.38, p = .028$, Cohen's $d = -0.53$, 95% CI [−9.72, −0.63], the difference between postrating and prerating scores were significantly different from zero, but not in the other two conditions: for incorrect feedback & gaze-follower, $t(19) = -0.77, p = .45$, Cohen's $d = -0.17$, 95% CI [−4.57, 2.12], and incorrect feedback & gaze-unfollower, $t(19) = 0.53, p = .61$, Cohen's $d = 0.12$, 95% CI [−3.06, 5.11]. The two-way repeated-measures ANOVA on differential trust scores revealed a significant interactive effect, $F(1, 19) = 14.18, p = .001, \eta_p^2 = 0.43$, but no significant main effect: for feedback, $F(1, 19) = 0.03, p = .86, \eta_p^2 = 0.002$, and for gaze-following, $F(1, 19) = 2.88, p = .11, \eta_p^2 = 0.13$. Post hoc analysis revealed the same pattern as in Experiment 1, meaning that when participants received correct feedback, they would raise trust ratings for the gaze-followers more than for gaze-unfollowers, $t(19) = 3.03, p = .007$, Cohen's $d = 0.96$, 95% CI [3.29, 17.96]. Nevertheless, this difference disappeared when they received incorrect feedback, $t(19) = -0.94, p = .36$, Cohen's $d = -0.20$, 95% CI [−7.27, 2.77] (see Fig. 5).

As for the differential invest scores, a repeated-measures ANOVA revealed a significant interaction, $F(1, 19) = 13.74, p = .001, \eta_p^2 = 0.42$, but no significant main effect: for feedback, $F(1, 19) = 0.88, p = .36, \eta_p^2 = 0.04$, and for gaze-following: $F(1, 19) = 4.12, p = .06, \eta_p^2 = 0.18$. Post hoc *t* tests revealed that people prefer to invest more money with gaze-followers than for gaze-unfollowers when they got correct feedback, $t(19) = 3.49, p = .002$, Cohen's $d = 1.21$, 95% CI [5.13, 20.52], but this effect disappeared in the incorrect feedback conditions, $t(19) = -1.20, p = .24$, Cohen's $d = -0.30$, 95% CI [−8.77, 2.37] (see Fig. 6).

¹ One-sample *t* tests results for four conditions: correct-feedback & gaze-follower, $t(19) = 1.32, p = .20$, Cohen's $d = 0.30$, 95% CI [−0.67, 2.97]; correct-feedback & gaze-unfollower, $t(19) = 0.95, p = .35$, Cohen's $d = 0.21$, 95% CI [−0.87, 2.32]; incorrect-feedback & gaze-follower, $t(19) = -0.99, p = .34$, Cohen's $d = -0.22$, 95% CI [−2.42, 0.87]; incorrect-feedback & gaze-unfollower, $t(19) = -1.18, p = .25$, Cohen's $d = -0.26$, 95% CI [−2.09, 0.59]; all *p* values reported were with Bonferroni correction.

Similarly, results for the standardized differential invest scores revealed a significant interaction between the two factors, $F(1, 19) = 10.24$, $p = .005$, $\eta_p^2 = 0.35$, but no significant main effect: for feedback, $F(1, 19) = 0.08$, $p = .78$, $\eta_p^2 = 0.004$, and for gaze-following, $F(1, 19) = 0.90$, $p = .35$, $\eta_p^2 = 0.05$. Post hoc test suggested there was a significant difference in the correct feedback conditions, $t(19) = 2.52$, $p = .02$, Cohen's $d = 0.83$, 95% CI [0.14, 1.51], but not in incorrect feedback conditions, $t(19) = -1.56$, $p = .14$, Cohen's $d = -0.40$, 95% CI [-0.94, 0.14] (see Fig. 7).

Discussion

Experiment 2 replicated our findings in Experiment 1; furthermore, the results also revealed the trust change induced by self-initiated social interactions influenced people's actual investment behaviors in economic games. These results demonstrated how social interaction influences trust evaluation from both psychological and behavioral levels. In this experiment, we showed participants the partners' repayment immediately after they made their investment to each partner in the Trust Game. However, an alternative view may argue that the previous partner's repayment might influence participant's decision in the next trial. For example, if a previous partner kept all the money to him, participants might feel deceived, and thus reduce the amount of investment to the next person. To rule out this possible confounding factor, we ran a supplementary experiment (Experiment S1) by presenting all the partners' repayment at the end of the whole experiment. The main results were replicated, confirming the reliability of our findings (see [Supplementary](#)).

Fig. 5 Differential trust scores in Experiment 2. Error bars indicate 95% CI within subjects (Cousineau, 2005). Asterisks indicate significant differences compared with zero (one-sample t test) and between groups (repeated-measures ANOVA). * $p < .05$. ** $p < .01$

Fig. 6 Differential invest scores in Experiment 2. Error bars indicate 95% CI within subjects (Cousineau, 2005). Asterisks indicate significant differences between groups (repeated-measures ANOVA). ** $p < .01$

General discussion

Similar to many other social judgments, our trust to other people is never a fixed impression but rather dynamically changing during social interactions (Tuk et al., 2009). In this current study, we investigated how noneconomic social interactions change interpersonal trust evaluations. Both studies show that other's gaze-following behavior influences people's trust evaluations, and these effects are modulated by affect states induced by feedback. Specifically, participants increased their trust ratings for gaze-followers and reduced that for gaze-unfollowers when these faces were always accompanied with positive feedback. However, these trust changes disappeared when people received negative feedback. Experiments 2 and S1 replicated this interesting change

Fig. 7 Standardized differential investment scores in Experiment 2. Error bars indicate 95% CI within subjects (Cousineau, 2005). Asterisks indicate significant differences between groups (repeated-measures ANOVA). * $p < .05$

pattern in the Trust Game, suggesting social interaction conveyed by gaze affects both the psychological and behavioral levels. Thereby, gaze-following behavior and feedback synergistically work on interpersonal trust change to partners, promoting the theoretical progress in interpersonal trust. That is, gaze-following behavior enhances interpersonal trust most effectively when the truster's affective state is positive.

In this way, our study discovered a dissociable influence of positive and negative affect on interpersonal trust. When faces were accompanied with positive feedback, participants increased their trust to those gaze-followers and decreased that for gaze-unfollowers. However, these trust changes disappeared in negative feedback condition, which indicated the important role of affective state in trust evaluation. Such results are similar to previous findings on how trust changes when people are in positive and negative moods. For example, researchers found that when primed positive affect, participants were likely to rate out-group members to be less trustworthy (Lount, 2010) and less willing to cooperate with them (Hertel, Neuhof, Theuer, & Kerr, 2000). However, participants' trust ratings were usually less likely to change in negative affective contexts (Bayliss et al., 2009) or for people with negative personalities (like depression; Bayliss, Tipper, Wakeley, Cowen, & Rogers, 2017). Furthermore, according to Fiedler's accommodation-assimilation model (Fiedler, 2001; Fiedler et al., 2003), positive feedback encouraged people to generate evaluations based upon heuristic processes, thus others' gaze-following behaviors would easily influence participants' trust ratings on them. Nevertheless, negative feedback caused people to considerate current situations more thoroughly and give more neutral judgments. The underlying reason might be the conflicting information participants received in the negative feedback conditions: agents who didn't follow us made the correct decision, while those who followed us turned out to be wrong. Thus, during these conflicting conditions with conflicting affective states, people preferred to choose more conservative judgement: remaining their first trust impression instead of changing it (Slovic, Finucane, Peters, & MacGregor, 2007).

Results from the current study also suggest corresponding behavioral changes when people change their trust ratings, which is in line with the findings in economic studies (e.g., van't Wout & Sanfey, 2008; J. P. Wilson, & Rule, 2015). Faces received higher ratings on their trustworthiness also received more investments from the participants. Besides, we also replicated these results after eliminating the potential interference of the partner's feedback in the Trust Game (Experiment S1). This suggests the trust impression formed outside of economic life can also bring difference to participants' economic behaviors, indicating the importance of trust evaluations on decision-making (Todorov, 2008). Considering that trust formulation in our study is under the mutual influence between social cues and decision

consequences, and changes in trust ratings can result in different behaviors in distinct contexts, future studies may further investigate the integrative model of trust (i.e., how people merge different sources of information and generalize the analytic product to various tasks).

Interestingly, we accidentally found a memory bias of feedback when asking participants to recall their initial trust ratings to the faces in Experiment 2. That is, people preferred to recall the faces associated with negative feedback to be less trustworthy than those with positive feedback. However, we didn't replicate this effect in Experiment S1 with the same procedure. In addition, when combining the data between these two experiments, the memory bias also disappeared, indicating that this might be a false positive effect found in Experiment 2. This reminds us the importance of replication. Furthermore, we did replicate our main results throughout the three experiments. Both Experiments 2 and S1 replicated the interactive effect of trust ratings, and Experiment S1 also replicated the same pattern of economic behavior in Experiment 2. All in all, our current studies provide a strong support to the interactive effect between gaze-following behavior and feedback on the interpersonal trust.

Our findings again confirmed the role of gaze, a rudimentary social signal, in shaping our interpersonal impressions. Many studies have revealed the gaze-cue effect in social cognition (Frischen et al., 2007). People regarded agents with more valid gaze cues as more attractive and trustworthy, on the contrary, those who always gaze towards invalid positions would evoke participant's negative emotional responses (e.g., Bayliss & Tipper, 2006; Ewing, Rhodes, & Pellicano, 2010; Rogers et al., 2014). Our findings further contributed to investigation of the relationship between gaze and social judgments, however, with a different form of gaze. It is different in two ways from the previous gaze-cue paradigm. First, in the gaze-following paradigm, participants become the active initiator of social interactive process, which enhance their sense of participation (Bayliss et al., 2013). Second, unlike gaze-cue that mainly serves as an informative cue, gaze-following acts more like a social glue, as it delivers our communicative intentions and supports cooperative behaviors (Farroni, Csibra, Simion, & Johnson, 2002). In a novel study, Farmer, Hertz, and Hamilton (2019) manipulated whether virtual agents would follow or unfollow participants' choice on selecting a more preferable painting. They found that agents who usually made the same choice as us would increase our judgments on their likeability and trustworthiness. Our results generalize these findings to gaze research and combine it with affective state to reveal more profound mechanisms underlying human interactions.

Our findings emphasized the importance of considering trust formation and change with social interactions. Our study is among the first few to reveal that gaze interaction directly affects our evaluation of other people's trustworthiness.

Compared with canonical cognitive research, the investigation on mechanism behind such routine social interactions is rather a new field. However, it contributes a great deal to interdisciplinary research, such as human–robot interactions. Future studies may focus on how other social signals, such as imitation, body postures, or biological motions, play a role in changing interpersonal trust.

Open practices statement Neither of the experiments reported in this article was formally preregistered. All data have been made publicly available via the Open Science Framework (<https://osf.io/an6w7/>).

Compliance with ethical standards

This research was supported by Humanities and Social Sciences Foundation of the Ministry of Education of China (19YJA190004) and Fundamental Research Funds for Colleges and Universities–Key Training Program for Young Teachers (19wkzd23).

Conflicts of interest All authors declare that they have no conflicts of interest.

References

- 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.
- Bayliss, A. P., Murphy, E., Naughtin, C. K., Kritikos, A., Schilbach, L., & Becker, S. I. (2013). “Gaze leading”: Initiating simulated joint attention influences eye movements and choice behavior. *Journal of Experimental Psychology: General*, 142(1), 76–92.
- Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments: Should I trust you? *Psychological Science*, 17(6), 514–520.
- Bayliss, A. P., Tipper, S. P., Wakeley, J., Cowen, P. J., & Rogers, R. D. (2017). Vulnerability to depression is associated with a failure to acquire implicit social appraisals. *Cognition and Emotion*, 31(4), 825–833.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1), 122–142.
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, 105(38), 14325–14329.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436.
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method. *Tutorials in Quantitative Methods for Psychology*, 1(1), 42–45.
- Driver, J., IV, Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, 6(5), 509–540.
- Edwards, S. G., Stephenson, L. J., Dalmaso, M., & Bayliss, A. P. (2015). Social orienting in gaze leading: A mechanism for shared attention. *Proceedings of the Royal Society B: Biological Sciences*, 282(1812), 20151141.
- 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.
- Farmer, H., Hertz, U., & Hamilton, A. (2019). *The neural basis of shared preference learning*. bioRxiv, 570762. <https://doi.org/10.1101/570762>
- Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, 99(14), 9602–9605.
- Fiedler, K. (2001). Affective states trigger processes of assimilation and accommodation. In L. L. Martin & G. L. Clore (Eds.), *Theories of mood and cognition: A user’s guidebook* (pp. 85–98). Mahwah, NJ: Erlbaum.
- Fiedler, K., Nickel, S., Asbeck, J., & Pagel, U. (2003). Mood and the generation effect. *Cognition and Emotion*, 17, 585–608.
- Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, 5(3), 490–495.
- 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.
- Hertel, G., Neuhof, J., Theuer, T., & Kerr, N. L. (2000). Mood effects on cooperation in small groups: Does positive mood simply lead to more cooperation? *Cognition and Emotion*, 14(4), 441–472.
- Hommel, B., & Colzato, L. S. (2015). Interpersonal trust: An event-based account. *Frontiers in Psychology*, 6, 1399.
- Krueger, F., & Meyer-Lindenberg, A. (2019). Toward a model of interpersonal trust drawn from neuroscience, psychology, and economics. *Trends in Neurosciences*, 42(2), 92–101.
- Kuzmanovic, B., Georgescu, A. L., Eickhoff, S. B., Shah, N. J., Bente, G., Fink, G. R., & Vogeley, K. (2009). Duration matters: Dissociating neural correlates of detection and evaluation of social gaze. *NeuroImage*, 46(4), 1154–1163.
- Lount, R. B., Jr. (2010). The impact of positive mood on trust in interpersonal and intergroup interactions. *Journal of Personality and Social Psychology*, 98(3), 420–433.
- Mason, M. F., Tatlow, E. P., & Macrae, C. N. (2005). The look of love: Gaze shifts and person perception. *Psychological Science*, 16(3), 236–239.
- Pashler, H., & Wagenmakers, E.-J. (2012). Editors’ introduction to the Special Section on Replicability in Psychological Science: A crisis of confidence? *Perspectives on Psychological Science*, 7(6), 528–530.
- R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Robbins, B. G. (2016). What is trust? A multidisciplinary review, critique, and synthesis. *Sociology Compass*, 10(10), 972–986.
- Rogers, R. D., Bayliss, A. P., Szepletowska, A., Dale, L., Reeder, L., Pizzamiglio, G., ... Tipper, S. P. (2014). I want to help you, but I am not sure why: Gaze-cuing induces altruistic giving. *Journal of Experimental Psychology: General*, 143(2), 763–777.
- Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology*, 104(3), 409–426.
- Schott, B. H., Minuzzi, L., Krebs, R. M., Elmenhorst, D., Lang, M., Winz, O. H., ... Düzel, E. (2008). Mesolimbic functional magnetic resonance imaging activations during reward anticipation correlate with reward-related ventral striatal dopamine release. *Journal of Neuroscience*, 28(52), 14311–14319.
- Schultz, W. (2007). Behavioral dopamine signals. *Trends in Neurosciences*, 30(5), 203–210.
- Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., & Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron*, 41(4), 653–662.

- Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research*, 177(3), 1333–1352.
- Steenbergen, H. V., Band, G. P., & Hommel, B. (2009). Reward counteracts conflict adaptation: Evidence for a role of affect in executive control. *Psychological Science*, 20(12), 1473–1477.
- Strachan, J. W., & Tipper, S. P. (2017). Examining the durability of incidentally learned trust from gaze cues. *The Quarterly Journal of Experimental Psychology*, 70(10), 2060–2075.
- Strachan, J. W. A., 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.
- Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*, 1124(1), 208–224.
- 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, 519–545.
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., ... Nelson, C. (2009). The NimStim Set of Facial Expressions: Judgments from untrained research participants. *Psychiatry Research*, 168(3), 242–249.
- Tuk, M. A., Verlegh, P. W., Smidts, A., & Wigboldus, D. H. (2009). Interpersonal relationships moderate the effect of faces on person judgments. *European Journal of Social Psychology*, 39(5), 757–767.
- van't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition*, 108(3), 796–803.
- Westfall, J. (2015). *PANGEA: Power analysis for general ANOVA designs*. Unpublished manuscript. Retrieved from <http://jakewestfall.org/publications/pangea.pdf>
- 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.
- Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, 26(8), 1325–1331.
- Wilson, R. K., & Eckel, C. C. (2006). Judging a book by its cover: Beauty and expectations in the trust game. *Political Research Quarterly*, 59(2), 189–202.

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