



Social overshadowing: Revisiting cue-competition in social interactions

Maiika Telga^{1,2,3} · José A. Alcalá^{4,5} · Cecilia Heyes⁶ · Gonzalo P. Urcelay⁵

Accepted: 26 November 2022
© The Author(s) 2023

Abstract

In a large variety of contexts, it is essential to use the available information to extract patterns and behave accordingly. When it comes to social interactions for instance, the information gathered about interaction partners across multiple encounters (e.g., trustworthiness) is crucial in guiding one's own behavior (e.g., approach the trustworthy and avoid the untrustworthy), a process akin to trial-by-trial learning. Building on associative learning and social cognition literatures, the present research adopts a domain-general approach to learning and explores whether the principles underlying associative learning also govern learning in social contexts. In particular, we examined whether overshadowing, a well-established cue-competition phenomenon, impacts learning of the cooperative behaviors of unfamiliar interaction partners. Across three experiments using an adaptation of the iterated Trust Game, we consistently observed a 'social overshadowing' effect, that is, a better learning about the cooperative tendencies of partners presented alone compared to those presented in a pair. This robust effect was not modulated by gender stereotypes or beliefs about the internal communication dynamics within a pair of partners. Drawing on these results, we argue that examining domain-general learning processes in social contexts is a useful approach to understanding human social cognition.

Keywords Overshadowing · Learning · Stereotypes · Cooperation · Trust

Being able to make accurate inferences about others and behave accordingly is key for smooth and harmonious relationships. To navigate social interactions successfully, one often monitors others' behavior across multiple encounters, infers their traits, and uses these inferences to decide how to act. For instance, repeated demonstrations of trustworthiness and cooperation over multiple encounters are likely responded to in a reciprocal fashion (King-Casas et al.,

2005). Building on previous work adopting a reinforcement learning approach (Chang et al., 2010; Cho & Hackel, 2022; Fareri et al., 2015), the present research examined whether the framework of associative learning may contribute to our understanding of trait inferences and social interactions.

Research in associative learning has demonstrated that across multiple trials, human and nonhuman animals learn the reward value associated with stimuli and adjust their behavior accordingly, approaching rewarding stimuli and avoiding nonrewarding stimuli (Bouton, 2007). A similar approach may account for human behaviors in social contexts: across repeated interactions, people may learn social values such as a person's generosity or trustworthiness, and based on this learning decide how to act whilst interacting with this person (Hackel et al., 2015). If learning of social values followed the same principles as learning in non-social contexts, the basic phenomena of associative learning would be expected in the social realm, as suggested by theories that underline the fundamental importance of domain-general, taxon-general processes in human cognition (Heyes, 2019; Lockwood et al., 2020; Reader, 2016; van Overwalle, 2011). In the present research, we test this hypothesis and examine

✉ Maiika Telga
myt1@st-andrews.ac.uk

✉ Gonzalo P. Urcelay
gonzalo.urcelay@nottingham.ac.uk

¹ School of Management, University of St Andrews, Gateway Building, North Haugh, St Andrews KY16 9AJ, UK

² University of Granada, Granada, Spain

³ University of Leicester, Leicester, UK

⁴ Complutense University of Madrid, Madrid, Spain

⁵ School of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, UK

⁶ University of Oxford, Oxford, UK

whether cue-competition, a hallmark of associative learning, also characterizes learning about the cooperative tendencies of unfamiliar interaction partners.

Cue-competition comprises a family of phenomena that underly learning about the relationships between events, for example a cue and an outcome, when such learning occurs in the presence of multiple cues. Among cue-competition phenomena, blocking and overshadowing have been investigated extensively. Blocking is the phenomenon in which learning about the reward value of a target cue X is impaired when trained in the presence of a cue A, which receives additional training either before (forward-blocking) or after (backward-blocking) XA compound training (Kamin, 1968; Shanks, 1985). Similarly, overshadowing refers to impaired learning about a cue X that was trained in compound with a cue B, compared to when trained alone (Pavlov, 1927). Blocking and overshadowing have been observed, not only in nonhuman animals, but in studies of human contingency judgment (Price & Yates, 1993; Shanks, 1985), Pavlovian conditioning (Martin & Levey, 1991), spatial learning (Herrera et al., 2022), or evaluative conditioning (Kattner & Green, 2015). All these examples suggest that cue-competition is due to domain-general mechanisms that are present in a wide range of species (Urcelay, 2017).

However, only a few studies have explored cue-competition in social contexts, and these have focused on blocking (Cramer et al., 1985; FeldmanHall et al., 2017; Mata et al., 2021). As far as we are aware, there have been only two reports of ‘social overshadowing’ in humans, dating back to the 1980s. Lanzetta and Orr (1980, 1981) found that social cues, such as faces varying in emotional expression, differentially impact learning about a tone-shock relationship in a fear conditioning paradigm: when a face expressing either fear or happiness and a neutral stimulus (i.e., tone) were trained in compound followed by a shock, fearful faces overshadowed the tone, and happy faces were overshadowed by the tone. Hence, Lanzetta and Orr found asymmetrical learning about the cues within the compound (i.e., the tone and the face), depending on the emotion portrayed – a reciprocal overshadowing effect suggesting that social information (faces) was being processed associatively with asocial information (tone). However, they did not include a control cue trained alone, leaving unanswered the question of whether learning about the target predictor (e.g., face) was overshadowed or potentiated by the presence of a second predictor (e.g., tone). In the current study, we aimed to overcome the limitations of previous overshadowing research in social contexts by examining trait inferences through social interactions in a classic overshadowing design with appropriate controls. Moreover, we aim to explore the potential interaction of learning processes with social mechanisms related to group membership such as stereotypes and intergroup biases.

To achieve this goal, we used an adaptation of the iterated Trust Game (King-Casas et al., 2005), which has been used extensively to explore trial-by-trial learning (e.g., Meidinger & Terracol, 2012), and the expression of social biases (e.g., Slonim & Guillen, 2010; Telga et al., 2018). The Trust Game recreates cooperation dynamics between people that resemble associative learning, as participants learn to approach rewarding stimuli (i.e., cooperative partners) and to avoid nonrewarding stimuli (i.e., noncooperative partners) on a trial-by-trial basis. Moreover, it allows us to explore potential social biases by manipulating partners’ features such as their social category membership. Across three experiments, we used the Trust Game to examine overshadowing in social interactions and its potential modulation by gender stereotypes (Experiments 1–3) and participants’ knowledge of the communication dynamics among partners (Experiment 3).

In Experiment 1, we used the Trust Game as a learning phase in which participants learned the cooperative (or uncooperative) tendencies of unfamiliar male and female game partners, presented either alone or in a pair (with a counterpart of the same gender). Next, in a test phase, participants were asked to indicate their likelihood of cooperating with each one of the partners from the Trust Game, now presented alone. We anticipated an overshadowing effect, that is, more learning of the cooperative tendencies of partners presented alone compared to those presented within a pair. Moreover, because women are associated with more cooperation than men (Buchan et al., 2008; Telga et al., 2018), we also assessed whether partners’ gender would impact the magnitude of the overshadowing effect, such that stereotype-inconsistent associations would be more vulnerable to competition, resulting in more overshadowing for male compared to female partners.

In Experiment 2, we used a similar design but introduced a new condition to explore whether the stereotypical belief that females are more trustworthy than males (Buchan et al., 2008) could modulate overshadowing. Like Lanzetta and Orr (1980, 1981), who observed that fearful faces overshadowed neutral stimuli in fear conditioning, we examined whether female faces would overshadow male faces when learning trustworthiness. For this, we introduced an additional mixed-gender pair (i.e., one male and one female partner within a pair) to the design.

Overshadowing is typically explained by associative processes (Mackintosh, 1976; Pearce, 1987) but the social overshadowing effects observed in Experiments 1 and 2 could instead have been due to reasoning (de Houwer et al., 2005). Participants may have learned less about their partner when the partner appeared in a pair because the participants inferred that the decision to cooperate or not was influenced by the other member of the pair, and therefore provided less information about the partner’s trustworthiness. This is a potential explanation because, when pairs of partners were

presented in Experiments 1 and 2, it was ambiguous whether the two members of the pair consensually decided how to act, or one member of the pair made a unilateral decision. We tested this alternative explanation in Experiment 3 by replicating Experiment 1 with an additional between-participants manipulation of instructions. In the non-ambiguous group, participants were instructed that partners presented in a pair made consensual decisions, hence their final decision reflected the cooperative tendency of both members. In the ambiguous group, ambiguity was similar to Experiments 1 and 2. We expected that if reasoning was responsible for the overshadowing effect, it should be replicated in the ambiguous group, but eliminated in the non-ambiguous group.

Methods

Participants

Sample size (a minimum of 40 participants per group) was based on previous research using the Trust Game (Telga et al., 2018). Participants were recruited through Prolific online crowdsourced platform and received a fixed financial compensation for their time (i.e., a minimum of £8.55/h, plus a bonus proportional to their performance (i.e., £0.011 per pound earned in the Trust Game), for a minimum of £2.85 and a maximum of £6 for the entire experiment. We recruited 48 participants in Experiment 1 (24 women, $M_{age} = 38.0$ years, $SD_{age} = 16.4$) and Experiment 2 (24 women, $M_{age} = 34.7$ years, $SD_{age} = 12.2$), and 96 participants (48 women, $M_{age} = 37.2$ years, $SD_{age} = 13.0$) in Experiment 3. A sensitivity analysis using G*Power (Faul et al., 2007) showed that with this sample, we could detect an effect of $\eta_p^2 = .15$ ($\beta = .80$, $\alpha = .05$) for the critical Partner Behavior \times Trial Type interaction reflecting an overshadowing effect. Note that in all experiments, the effect observed was larger than .15. To control for the impact of cultural background on cooperation decisions (Gächter et al., 2010), we pre-screened participants to select only those who were born, had spent most of their childhood, and were currently living in the same country – the United Kingdom. We also pre-screened participants who had already taken part in studies of this experimental series. Consent was obtained at the beginning of the experiments. All the studies were approved by the University of Leicester ethics committee (ref: 27997).

Stimuli and materials

Gorilla software (Anwyl-Irvine et al., 2020) was used for stimulus presentation and data collection. All the materials used in these experiments are available on the Gorilla Open Materials Repository (<https://app.gorilla.sc/openmaterials/451374>). Sixteen colored photographs were extracted from

the Chicago Face Database (Ma et al., 2015) to represent the game partners (six men and six women in Experiments 1 and 3, eight men and eight women in Experiment 2). All faces were presented against a white background with a direct gaze and a neutral emotional expression. The Chicago Face Database is a standardized set of faces including ratings about the perceived sex, race, or trustworthiness of all targets. This information was used to select the faces to be used in the Trust Game. Specifically, we ensured that all faces belonged to the same racial category (i.e., all White), and were approximately the same age ($M = 27.6$ years, $SD = 1.6$ in Experiments 1 and 3; $M = 27.1$ years, $SD = 1.5$ in Experiment 2). We also ensured that there was no difference between male and female partners regarding their age, $F(1,10) = .02$, $p = .88$, $\eta_p^2 < .01$, in Experiments 1 and 3, $F(1, 14) = .38$, $p = .55$, $\eta_p^2 = .03$, in Experiment 2, and their perceived trustworthiness, $F(1,10) = .60$, $p = .46$, $\eta_p^2 = .06$, in Experiments 1 and 3, $F(1, 14) = .12$, $p = .73$, $\eta_p^2 < .01$ in Experiment 2.

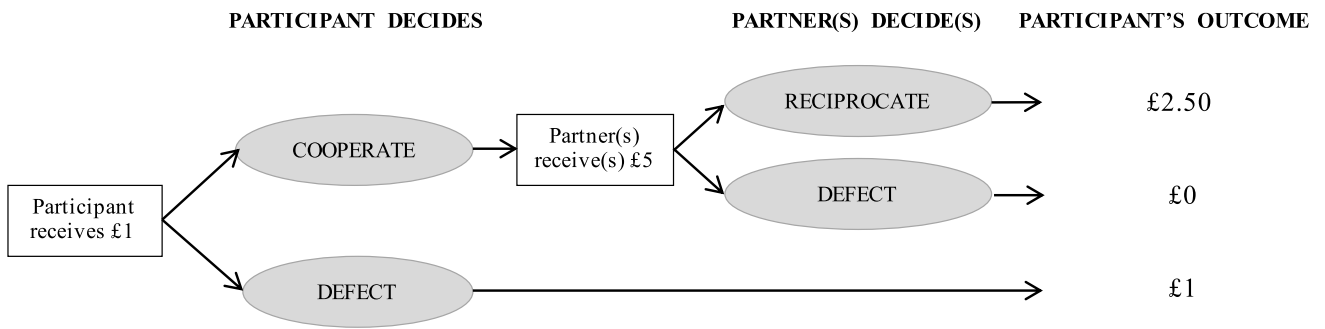
Procedure

After providing consent, participants started the experiment, which comprised three phases: baseline, learning, and test. The complete instructions provided to the participants are provided in Sections 1 and 2 of the Online Supplementary Materials (OSM).

First, the baseline allowed us to assess participants' spontaneous likelihood to cooperate with each one of the partners before the Trust Game. Each trial consisted of a single display in which the photograph of one partner was presented in the center of the screen and the question “*How likely are you to spontaneously cooperate with this person?*” was displayed below the photograph. To respond, participants used a horizontal slider ranging from 0 (Very Unlikely) to 100 (Very Likely). Once they had indicated their ratings, participants moved to the next trial by pressing the “*Continue*” button. Although there was no time limit, participants were encouraged to respond as fast as possible. The order of presentation of the faces was randomized for each participant.

Next, participants performed the learning phase, consisting of an adaptation of the Trust Game (Berg et al., 1995; Telga et al., 2018), as shown in Fig. 1. At the beginning of each trial, participants virtually received £1 and had to decide what to do with it. After a fixation point appeared in the center of the screen for 200 ms, participants saw the image of the partner or pair of partners of this round for 1,000 ms. Next, the question “*Do you cooperate?*” appeared below the photograph(s). Participants had 1,500 ms to indicate whether they cooperated by pressing the ‘1’ key or did not cooperate by pressing the ‘0’ key. If participants decided to cooperate, the partner or pair of partners

A. Pay-offs structure



B. Trial structure



Fig. 1 Pay-off structure (a) and timeline of one trial (b) in the Trust Game

in this round would ostensibly receive the initial amount multiplied by 5 (i.e., £5) and decide whether to either give £2.50 back to the participant (i.e., reciprocate) or to keep all the money (i.e., defect). In this case, they were presented a 2,000-ms feedback display informing of their partner(s) decision (i.e., either reciprocated or defected) together with their final monetary outcome (i.e., either £2.50 or £0). If, in contrast, participants decided not to cooperate, they would keep the initial £1 and the partner(s) would receive nothing. In that case, they were presented a 2,000-ms feedback screen informing of their monetary outcome (i.e., £1) as well as the decision that the partner would have made if they had cooperated (i.e., the partner could have either reciprocated or defected). Trials were separated by a 1,000-ms inter-trial interval (ITI).

Before they started the Trust Game, participants received detailed instructions about the task, specifying the payoff structure of a trial. Next, they performed an eight-trial practice with four partners (different from those presented in the Trust Game). For these practice trials, participants were encouraged to alternately cooperate and not cooperate, so they could become familiarized with the four different types of feedback. They were also told that their responses in the practice were not considered to compute their final earnings. Upon completion of the practice trials, participants started the Trust Game in which we manipulated their partners' gender, partners' behavior, and the type of trials.

In Experiment 1, we manipulated within-participants the cooperative behaviors of 12 different partners: six men and six women. These partners were presented either individually (single partner condition) or in a pair (pair of partners condition). Partners presented in a pair were of the same gender (i.e., either two men or two women) and always displayed with the same counterpart across trials. Whether partners presented in a pair were displayed on the left or on the right side of the screen was counterbalanced across trials. Within these two trial type conditions, half of the partners were always cooperative and allowed participants to earn money (i.e., cooperative partners condition) while the other half of partners were always noncooperative, leading participants to no monetary gain (i.e., noncooperative partners condition). The eight types of trials resulting from the factorial combination of partners' gender (female vs. male), trial type (single vs. pair), and partners' behavior (cooperative vs. noncooperative) were presented 20 times, resulting in 160 training trials.

Experiment 2 used a similar design, except that two mixed-gender pairs (one cooperative pair, and one noncooperative pair) were added to the aforementioned design. Therefore, participants were presented with 16 different partners, eight men and eight women, for a total of ten types of trials presented 20 times, resulting in 200 trials.

Finally, Experiment 3 used a similar design to Experiment 1, with an additional between-group manipulation of the instructions provided to participants. Specifically, in the non-ambiguous group, participants were told that when a pair of partners decided to cooperate (defect), both of them had agreed and therefore both were cooperative (noncooperative). In the ambiguous group, participants were told that when a pair of partners made a decision to cooperate (defect), it meant that either one of them or both of them made the final call, and therefore either one of them or both were cooperative (noncooperative). After the instructions, we introduced a manipulation check to ensure that participants knew how decisions were made within a pair. Participants who failed this check re-experienced the instructions until they correctly responded to all questions.

In all experiments, the order of presentation of the trials was pseudo-randomized within ten mini-blocks of two presentations of each type of trial. We also ensured that across participants, all the photographs of the same gender were alternately associated with the four experimental conditions resulting from the combination of trial type and partner behavior. Participants had two opportunities to take a short break, up to 1 min, after Block 3 and after Block 7.

The test phase was similar to the baseline phase, except that participants were asked to evaluate the likelihood that they would cooperate with each partner on the basis of what they had learned during the learning phase, rather than based on their first impression. The experiment took around 20 min to complete.

Data exclusion

Participants who responded on fewer than 20% of the trials in the learning phase were excluded from the analyses. Following this criterion, we eliminated two participants from the analyses of Experiment 2, leaving 46 participants. In Experiments 1 and 3, data from all participants were included in the analyses.

Design and analyses

The baseline allowed us to obtain an initial score of participants' likelihood to cooperate with each partner. This DV was analyzed as a function of partner gender (female vs. male), while controlling for participant gender (female vs. male). With this design, a cooperation bias toward one of the two genders should be reflected in a main effect of partner gender. We expected a higher likelihood to cooperate with female than with male partners (Buchan et al., 2008; Telga et al., 2018).

For the learning phase in Experiment 1, we analyzed participants' cooperation rate (i.e., the proportion of trials on which they cooperated with their partner(s)) as a function

of partner gender (female vs. male), partner behavior (cooperative vs. noncooperative), trial type (single vs. pair), and block (1–10). In Experiment 2, the same DV was analyzed as a function of partner behavior (cooperative vs. noncooperative), trial type (female single, male single, female pair, male pair, mixed-gender pair), and block (1–10). In Experiment 3, we used the same design as in Experiment 1, with an additional between-participants variable of ambiguity. In all experiments, we controlled for participants' gender by introducing this variable in the ANOVA. The analyses of the learning phase of all experiments are reported in Section 3 of the OSM.

For the test phase, we subtracted the likelihood to cooperate with each partner in the baseline phase from the likelihood to cooperate with them in the test phase, to control for potential spontaneous biases in cooperation. This DV was analyzed as a function of partner gender (male vs. female), partner behavior (cooperative vs. noncooperative), and trial type (single vs. same-gender pair (vs. mixed-gender pair in Experiment 2)). In Experiment 3, we added the between-participants variable ambiguity (ambiguous vs. non-ambiguous) to the aforementioned design. Overall, we expected an increase in the likelihood to cooperate with cooperative partners, and a decrease in the likelihood to cooperate with noncooperative partners. Critically, with this design, an overshadowing effect should be reflected in a significant Partner Behavior \times Trial Type interaction, indicating a larger increase in the likelihood to cooperate with cooperative partners presented alone compared to those presented in a pair, and a larger decrease in the likelihood to cooperate with noncooperative partners presented alone compared to those presented in a pair.

Results

Baseline phase

A perusal of Table 1 suggests that in all experiments, participants were consistently more likely to cooperate with female than with male partners. In fact, a mixed-design ANOVA with participant gender (female vs. male) as a between-participants factor and partner gender as a

within-participants variable revealed that the main effect of partner gender was significant in Experiment 1, $F(1, 46) = 8.18, p < .01, \eta_p^2 = .15, 90\% \text{ CI } [.03, .30]$, Experiment 2, $F(1, 44) = 9.06, p < .01, \eta_p^2 = .17, 90\% \text{ CI } [.03, .33]$, and Experiment 3, $F(1, 94) = 7.24, p < .01, \eta_p^2 = .07, 90\% \text{ CI } [.01, .17]$. These effects were not qualified by participant gender in Experiment 1, $F(1, 46) = 2.22, p = .14, \eta_p^2 = .05, 90\% \text{ CI } [.00, .17]$, Experiment 2, $F(1, 44) = 0.16, p = .70, \eta_p^2 < .01, 90\% \text{ CI } [.00, .08]$, or Experiment 3, $F(1, 94) = 0.07, p = .79, \eta_p^2 < .01, 90\% \text{ CI } [.00, .03]$.

Test phase

In all experiments, after interacting with cooperative and non-cooperative game partners, participants increased their likelihood to cooperate with cooperative partners, and decreased their likelihood to cooperate with non-cooperative partners. Critically, and as shown in Fig. 2, the tendency was more pronounced for partners presented alone compared to those presented in a same-gender or mixed-gender pair, consistent with the expected overshadowing effect. A mixed-design ANOVA with partner gender, partner behavior, and trial type as within-participants variables, and participant gender (and ambiguity in Experiment 3) as between-participants factors revealed that the critical Partner Behavior \times Trial Type interaction was significant in Experiment 1, $F(1, 46) = 33.43, p < .01, \eta_p^2 = .42, 90\% \text{ CI } [.23, .55]$, Experiment 2, $F(2, 88) = 24.32, p < .01, \eta_p^2 = .36, 90\% \text{ CI } [.22, .46]$, and Experiment 3, $F(1, 92) = 64.61, p < .01, \eta_p^2 = .41, 90\% \text{ CI } [.28, .51]$. Moreover, this interaction was not qualified by partner gender or participant gender in Experiments 1–3, nor by ambiguity in Experiment 3.¹ Inferential and Bayesian statistics of the nonsignificant interactions are provided in the OSM (Table 1). These revealed that overall, the data did not fit well with models including these interactions. We, therefore, analyzed the main effect of trial type in the cooperative and noncooperative partners conditions for each experiment.

In Experiment 1, and as shown in Fig. 2a, participants were more likely to cooperate with cooperative partners presented alone compared to those presented in a pair. Similarly, their likelihood to cooperate with noncooperative

Table 1 Descriptive statistics for the baseline phase

	Female partners	Male partners
Experiment 1	55.9 (12.7)	50.5 (11.3)
Experiment 2	53.8 (9.6)	49.0 (10.2)
Experiment 3	54.6 (12.3)	50.8 (11.2)

Means and standard deviations (in parentheses) of the likelihood to cooperate with female and male partners in the three experiments

¹ Note that to ensure that the absence of gender differences was not due to low statistical power, we analyzed the three experiments together (excluding the mixed-gender pair condition in Experiment 2). With this larger sample ($N = 190$), we observed the same pattern as observed in each experiment alone, that is, an overshadowing effect reflected in a significant Partner Behavior \times Trial Type interaction, $F(1, 184) = 133.83, p < .01, \eta_p^2 = .42$, not qualified by Participant Gender, $F(1, 184) = 0.89, p = .35, \eta_p^2 < .01$, nor by any other variables, largest $F(1, 184) = 1.39, p = .25, \eta_p^2 = .02$, for the Partner Behavior \times Trial Type \times Partner Gender \times Experiment interaction.

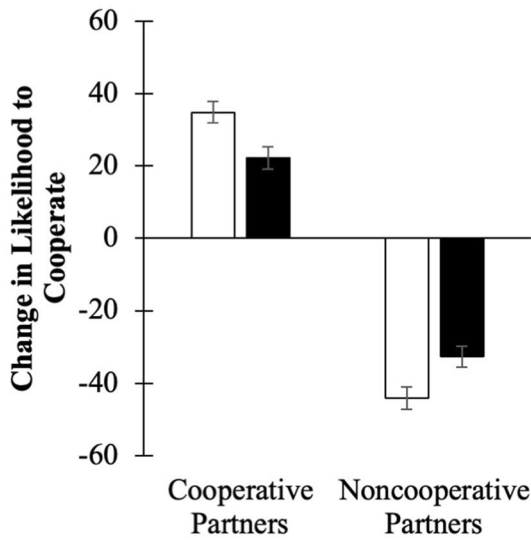
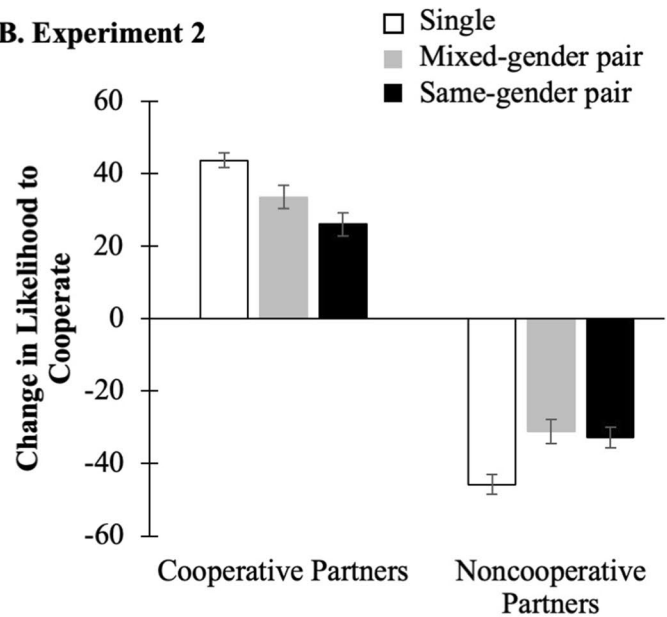
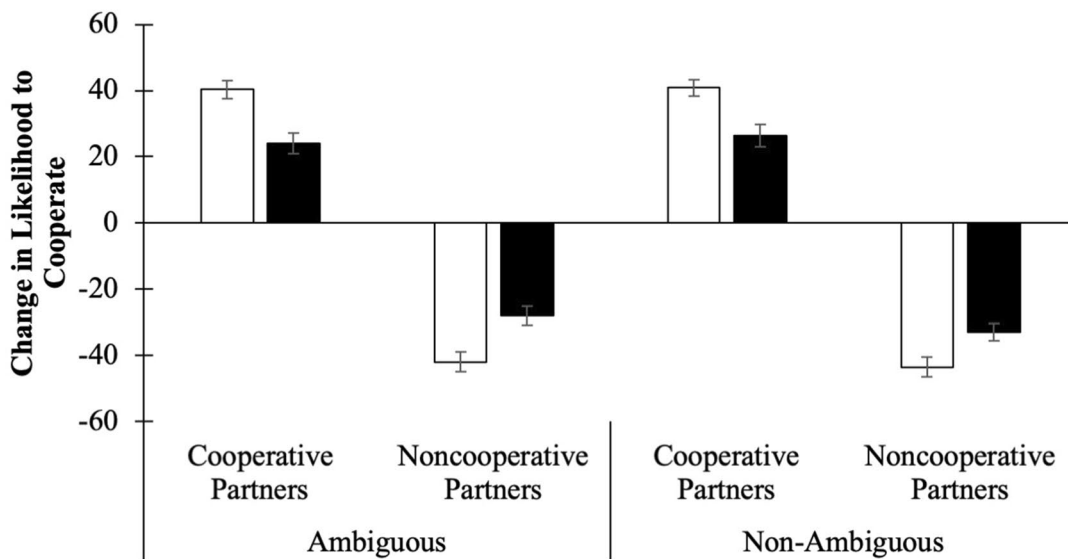
A. Experiment 1**B. Experiment 2****C. Experiment 3**

Fig. 2 Change in the likelihood to cooperate from the baseline to the test phase in Experiment 1 (a), Experiment 2 (b), and Experiment 3 (c). *Note:* Error bars represent the standard error of the mean with within-participant correction following Cousineau and O'Brien (2014)

partners decreased more towards partners presented alone compared to those presented in a pair. In fact, the main effect of Trial Type was significant in both the cooperative, $F(1, 47) = 19.06, p < .01, \eta_p^2 = .29, 90\% \text{ CI } [.12, .44]$, and the noncooperative, $F(1, 47) = 17.81, p < .01, \eta_p^2 = .28, 90\% \text{ CI } [.11, .42]$, partners conditions.

In Experiment 2, in which we introduced a mixed-gender pair of partners, a similar pattern was observed. As shown in Fig. 2b, participants showed a larger increase in

their likelihood to cooperate with cooperative partners presented alone compared to those presented in same-gender or mixed-gender pairs. The main effect of trial type was, in fact, significant in the cooperative partners condition, $F(2, 90) = 15.12, p < .01, \eta_p^2 = .25, 90\% \text{ CI } [.12, .36]$, revealing significant differences in the increase in the likelihood to cooperate between the single and the same-gender conditions, $F(1, 45) = 46.90, p < .01, \eta_p^2 = .51, 90\% \text{ CI } [.33, .62]$, and the single and the mixed-gender

conditions, $F(1, 45) = 10.71, p < .01, \eta_p^2 = .19, 90\% \text{ CI } [.05, .35]$, but not between same-gender and mixed-gender conditions, $F(1, 45) = 3.78, p = .06, \eta_p^2 = .08, 90\% \text{ CI } [.00, .22]$. Moreover, within the mixed-gender pair, the likelihood to cooperate with the female and the male partner did not significantly differ, $F(1, 45) = 0.08, p = .79, \eta_p^2 < .01, 90\% \text{ CI } [.00, .06]$, with a $\text{BF}_{01} = 4.56$ supporting the absence of differences. As depicted in Fig. 2b, participants showed a similar pattern of behavior towards noncooperative partners. In fact, the main effect of trial type was also significant in the noncooperative partners condition, $F(2, 90) = 11.03, p < .001, \eta_p^2 = .20, 90\% \text{ CI } [.08, .30]$, with significant differences between the single and the same-gender conditions, $F(1, 45) = 17.90, p < .01, \eta_p^2 = .29, 90\% \text{ CI } [.11, .43]$, the single and the mixed-gender conditions, $F(1, 45) = 20.27, p < .01, \eta_p^2 = .31, 90\% \text{ CI } [.13, .46]$, but not between the same-gender and the mixed-gender conditions, $F(1, 45) = 0.18, p = .68, \eta_p^2 < .01, 90\% \text{ CI } [.00, .08]$. Again, the decrease in likelihood to cooperate with the female or the male partners within the mixed-gender pair did not differ, $F(1, 45) = 0.25, p = .62, \eta_p^2 < .01, 90\% \text{ CI } [.00, .09]$, with a $\text{BF}_{01} = 4.10$ supporting the absence of differences.

Finally, in Experiment 3, in which we manipulated between participants that decision-making within a pair was either ambiguous or non-ambiguous, a similar pattern was observed. As depicted in Fig. 2c, participants from the two groups showed a larger increase in their likelihood to cooperate with cooperative partners presented alone relative to cooperative partners presented in a pair. The same was observed in the decrease for non-cooperative partners, regardless of the ambiguity on how partners in a pair made the decision to either cooperate or not cooperate. The main effect of trial type was significant in both the cooperative, $F(1, 95) = 46.54, p < .01, \eta_p^2 = .33, 90\% \text{ CI } [.20, .43]$, and the noncooperative, $F(1, 95) = 34.66, p < .01, \eta_p^2 = .27, 90\% \text{ CI } [.15, .38]$, partners conditions.

Discussion

Across three experiments, we investigated whether overshadowing characterizes learning about the cooperative tendencies of unfamiliar game partners. We found that participants learned more about the cooperative tendencies of partners presented alone compared to those presented in a pair, a result that was consistent across all experiments, and unaffected by gender category (Experiments 1, 2, and 3) and instructions (Experiment 3) manipulations.

The results of Experiment 3 suggest that the social overshadowing effects were due to associative processes rather than reasoning. If participants had inferred that a partner's

behavior provides less information about their cooperative tendencies when the partner appeared in a pair, because the other member of the pair may have influenced the partner's decision, the overshadowing effect should have disappeared in the non-ambiguous group. In fact, there was no difference between the non-ambiguous and the ambiguous groups. Across the three experiments, the overshadowing effect appeared in all eight conditions: when the partner was cooperative and non-cooperative, in a same- and mixed-gender pair, with ambiguous and non-ambiguous instructions. Although gender did not seem to interact with the overshadowing effect, we did observe a reliance on gender stereotypes at baseline with more cooperation with female than with male partners (Buchan et al., 2008; Telga & Lupiáñez, 2021), suggesting that this paradigm is sensitive to gender-related biases. Thus, the results of the present study suggest that models of associative learning (e.g., Rescorla & Wagner, 1972), developed in the nonsocial domain and predominantly in research with nonhuman animals, may advance our understanding of human social cognition (Behrens et al., 2008; Heyes, 2012), including the processes underlying cooperation and trust (FeldmanHall & Dunsmoor, 2018).

A priority for future research is to work out how domain-general processes such as associative learning and reasoning produce and interact with specifically social processes (Heyes, 2019; Heyes et al., 2020). Consistent with our results, a recent report has shown that social group membership does not modulate blocking in social learning (Vaz et al., 2022). However, there is also evidence that learning processes can interact with mechanisms specific to the social world. For instance, it has been shown that fear extinction, one of the most elemental associative learning phenomena, is modulated by stereotypical beliefs (Navarrete et al., 2009; Olsson et al., 2005). Similarly, using economic games, FeldmanHall et al. (2017) observed blocking in social decision making, but only in the gain rather than the loss domain, suggesting that specific social expectations may impact people's reasoning about the likelihood that an event will occur (i.e., what are the odds for a person to steal or be altruistic). Moreover, the impact of social category information (e.g., gender stereotypes) largely depends on the salience of such categories (Freeman & Ambady, 2011) and the cognitive demands of the task (Gilbert & Hixon, 1991). This raises the possibility that more salient social categories, or a more complex learning task, would promote a larger reliance on stereotypical trait attribution, as observed in previous research (Telga et al., 2018). Future studies should aim to understand under which circumstances specific social mechanisms interact with domain-general processes to guide our behavior.

One of the challenges of studying learning processes in social contexts rests in recreating ecologically valid scenarios while maintaining methodological rigor. On that note, our

adaptation of the Trust Game deviated from most real-life situations in that participants received feedback about their partners' trustworthiness when they chose both to trust and not to trust them. This ensured that all participants started the test phase with the same exposure to their partners' cooperative behaviors, so any observed effects are better explained by our experimental manipulations than by participants' varying knowledge about their partners' tendencies to cooperate. However, this may have resulted in a more accurate learning of others' trustworthiness compared to what would be expected with an asymmetrical feedback design (Fetchenhauer & Dunning, 2010). Further research is needed to determine whether, with an asymmetrical feedback design and therefore poorer learning, the degraded contingency would attenuate the overshadowing effect (Urceley, 2017).

The extension of research on cue-competition phenomena to social contexts opens interesting research avenues for both learning and social psychology theorists. The large body of research generated by the discovery of blocking and overshadowing has allowed researchers to propose prominent models that account for not only the phenomena themselves but also their neural substrates (Waelti et al., 2001), and their modulation by different factors (e.g., physical salience; Mackintosh, 1976). Thus, exploring cue-competition in social contexts offers to learning researchers a new area to test these theories and assess their generality and validity in the domain of complex social interactions. For instance, would social psychological salience impact overshadowing in a similar way as physical salience? Moreover, it offers a new perspective for social psychologists to explore phenomena such as impression formation and trait inferences. For instance, is the dynamism and malleability of person construal (Freeman & Ambady, 2011) subject to retrospective re-evaluation, as observed in backward blocking (Shanks, 1985)? A convergent approach to these phenomena, drawing on learning theory and social psychology, promises to provide a more comprehensive view of human social behavior.

Overall, the present research provides novel evidence that cue-competition may account for social behaviors such as trait inferences, trust, and cooperation. More broadly, it suggests that models of associative learning could be an asset in research on human social cognition.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13423-022-02229-3>.

Funding This research was supported by an Economic and Social Research Council grant (ES/R011494/2) awarded to Gonzalo P. Urceley and conducted while Maïka Telga was visiting the University of Leicester from the University of Granada as an Honorary Fellow. Note that Maïka Telga is now working at the University of St Andrews. Part of these findings were presented by Maïka Telga at the Virtual Meeting of the Spanish Comparative Psychology Association (2021) and at the Eastern Psychological Association Conference (2022).

Declarations

Conflicts of interest The authors declare no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. <https://doi.org/10.3758/s13428-019-01237-x>
- Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, *456*(7219), 245–249. <https://doi.org/10.1038/nature07538>
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*(1), 122–142. <https://doi.org/10.1006/game.1995.1027>
- Bouton, M. E. (2007). Learning and behavior: A contemporary synthesis. Sinauer Associates.
- Buchan, N. R., Croson, R., & Solnick, S. (2008). Trust and gender: An examination of behavior and beliefs in the Investment Game. *Journal of Economic Behavior & Organization*, *68*(3), 466–476. <https://doi.org/10.1016/j.jebo.2007.10.006>
- Chang, L. J., Doll, B. B., van't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, *61*(2), 87–105. <https://doi.org/10.1016/j.cogpsych.2010.03.001>
- Cho, H. J., & Hackel, L. M. (2022). Instrumental learning of social affiliation through outcome and intention. *Journal of Experimental Psychology: General*, *151*(9), 2204–2221. <https://doi.org/10.1037/xge0001190>
- Cousineau, D., & O'Brien, F. (2014). Error bars in within-subject designs: A comment on Baguley (2012). *Behavior Research Methods*, *46*(4), 1149–1151. <https://doi.org/10.3758/s13428-013-0441-z>
- Cramer, R. E., Weiss, R. F., Steigleder, M. K., & Balling, S. S. (1985). Attraction in context: Acquisition and blocking of person-directed action. *Journal of Personality and Social Psychology*, *49*(5), 1221–1230. <https://doi.org/10.1037/0022-3514.49.5.1221>
- de Houwer, J., Vandorpe, S., & Beckers, T. (2005). Evidence for the role of higher order reasoning processes in cue competition and other learning phenomena. *Learning & Behavior*, *33*(2), 239–249. <https://doi.org/10.3758/BF03196066>
- Fareri, D. S., Chang, L. J., & Delgado, M. R. (2015). Computational Substrates of Social Value in Interpersonal Collaboration. *The Journal of Neuroscience*, *35*(21), 8170. <https://doi.org/10.1523/JNEUROSCI.4775-14.2015>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. <https://doi.org/10.3758/bf03193146>

- FeldmanHall, O., & Dunsmoor, J. E. (2018). Viewing adaptive social choice through the lens of associative learning. *Perspectives on Psychological Science*, *14*(2), 175–196. <https://doi.org/10.1177/1745691618792261>
- FeldmanHall, O., Dunsmoor, J. E., Kroes, M. C. W., Lackovic, S., & Phelps, E. A. (2017). Associative learning of social value in dynamic groups. *Psychological Science*, *28*(8), 1160–1170. <https://doi.org/10.1177/0956797617706394>
- Fetchenhauer, D., & Dunning, D. (2010). Why so cynical?: Asymmetric feedback underlies misguided skepticism regarding the trustworthiness of others. *Psychological Science*, *21*(2), 189–193. <https://doi.org/10.1177/0956797609358586>
- Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, *118*(2), 247–279. <https://doi.org/10.1037/a0022327>
- Gächter, S., Herrmann, B., & Thöni, C. (2010). Culture and cooperation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1553), 2651–2661. <https://doi.org/10.1098/rstb.2010.0135>
- Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, *60*(4), 509–517. <https://doi.org/10.1037/0022-3514.60.4.509>
- Hackel, L. M., Doll, B. B., & Amodio, D. M. (2015). Instrumental learning of traits versus rewards: Dissociable neural correlates and effects on choice. *Nature Neuroscience*, *18*(9), 1233–1235. <https://doi.org/10.1038/nn.4080>
- Herrera, E., Alcalá, J. A., Tazumi, T., Buckley, M. G., Prados, J., & Urcelay, G. P. (2022). Temporal and spatial contiguity are necessary for competition between events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(3), 321–347. <https://doi.org/10.1037/xlm0001108>
- Heyes, C. (2012). What's social about social learning? *Journal of Comparative Psychology*, *126*(2), 193–202. <https://doi.org/10.1037/a0025180>
- Heyes, C. (2019). Précis of cognitive gadgets: The cultural evolution of thinking. *Behavioral and Brain Sciences*, *Pinker, 1994*. <https://doi.org/10.1017/S0140525X18002145>
- Heyes, C., Chater, N., & Dwyer, D. M. (2020). Sinking in: The peripheral Baldwinisation of human cognition. *Trends in Cognitive Sciences*, *24*(11), 884–899. <https://doi.org/10.1016/j.tics.2020.08.006>
- Kamin, L. J. (1968). Attention-like processes in classical conditioning. In M. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9–31). Miami, FL: University of Miami Press.
- Kattner, F., & Green, C. S. (2015). Cue competition in evaluative conditioning as a function of the learning process. *Acta Psychologica*, *162*, 40–50. <https://doi.org/10.1016/j.actpsy.2015.09.013>
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, *308*(5718), 78–83. <https://doi.org/10.1126/science.1108062>
- Lanzetta, J. T., & Orr, S. P. (1980). Influence of facial expressions on the classical conditioning of fear. *Journal of Personality and Social Psychology*, *39*(6), 1081–1087. <https://doi.org/10.1037/h0077713>
- Lanzetta, J. T., & Orr, S. P. (1981). Stimulus properties of facial expressions and their influence on the classical conditioning of fear. *Motivation and Emotion*, *5*(3), 225–234. <https://doi.org/10.1007/BF00993886>
- Lockwood, P. L., Apps, M. A. J., & Chang, S. W. C. (2020). Is there a 'social' brain? Implementations and algorithms. *Trends in Cognitive Sciences*, *24*(10), 802–813. <https://doi.org/10.1016/j.tics.2020.06.011>
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavioral Research Methods*, *47*(4), 1122–1135. <https://doi.org/10.3758/s13428-014-0532-5>
- Mackintosh, N. J. (1976). Overshadowing and stimulus intensity. *Animal Learning & Behavior*, *4*(2), 186–192. <https://doi.org/10.3758/BF03214033>
- Martin, I., & Levey, A. B. (1991). Blocking observed in human eyelid conditioning. *Quarterly Journal of Experimental Psychology*, *43*, 233–255. <https://doi.org/10.1007/BF02691035>
- Mata, A., Vaz, A., & Palma, T. A. (2021). Cue competition in mental state inference: Blocking effects in learning to interpret the behaviors of others. *Journal of Experimental Social Psychology*, *95*, 104122. <https://doi.org/10.1016/j.jesp.2021.104122>
- Meidinger, C., & Terracol, A. (2012). Learning in the trust game. *Revue de Philosophie Economique*, *13*(1), 155–174. <https://doi.org/10.3917/rpec.131.0155>
- Navarrete, C. D., Olsson, A., Ho, A. K., Mendes, W. B., Thomsen, L., & Sidanius, J. (2009). Fear extinction to an out-group face: The role of target gender. *Psychological Science*, *20*(2), 155–158. <https://doi.org/10.1111/j.1467-9280.2009.02273.x>
- Olsson, A., Ebert, J. P., Banaji, M. R., & Phelps, E. A. (2005). The role of social groups in the persistence of learned fear. *Science*, *309*(5735), 785–787. <https://doi.org/10.1126/science.1113551>
- Pavlov, I. P. (1927). *Conditioned reflexes*. (G. V. Anrep, Ed & Trans.). Oxford University Press.
- Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, *94*(1), 61–73. <https://doi.org/10.1037/0033-295X.94.1.61>
- Price, P. C., & Yates, J. F. (1993). Judgmental overshadowing: Further evidence of cue interaction in contingency judgment. *Memory & Cognition*, *21*(5), 561–572. <https://doi.org/10.3758/bf03197189>
- Reader, S. M. (2016). Animal social learning: Associations and adaptations. *F1000Research*, *5*, F1000 Faculty Rev—2120. <https://doi.org/10.12688/f1000research.7922.1>
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.
- Shanks, D. R. (1985). Forward and Backward Blocking in Human Contingency Judgement. *The Quarterly Journal of Experimental Psychology Section B*, *37*(1b), 1–21. <https://doi.org/10.1080/14640748508402082>
- Slonim, R., & Guillen, P. (2010). Gender selection discrimination: Evidence from a Trust game. *Journal of Economic Behavior & Organization*, *76*(2), 385–405. <https://doi.org/10.1016/j.jebo.2010.06.016>
- Telga, M., & Lupiáñez, J. (2021). Older and younger adults perform similarly in an iterated Trust Game. *Frontiers in Psychology*, *12*(747187). <https://doi.org/10.3389/fpsyg.2021.747187>
- Telga, M., de Lemus, S., Cañadas, E., Rodríguez-Bailón, R., & Lupiáñez, J. (2018). Category based learning about deviant outgroup members hinders performance in trust decision making. *Frontiers in Psychology*, *9*(1008). <https://doi.org/10.3389/fpsyg.2018.01008>
- Urcelay, G. P. (2017). Competition and facilitation in compound conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*(4), 303–314. <https://doi.org/10.1037/xan0000149>
- van Overwalle, F. (2011). Social learning and connectionism. In T. R. Schachtman, and S. S. Reilly (Eds.), *Associative Learning and Conditioning Theory: Human and Non-Human Applications* (pp. 345–375). Oxford Academic. <https://doi.org/10.1093/acprof:oso/9780199735969.003.0104>

- Vaz, A., Mata, A., & Palma, T. A. (2022). Blocking effects in social inference generalize across targets: Learning to interpret a person's behavioral cues interferes with learning about other people's behavior. *Social Cognition*, 40(4), 364–386. <https://doi.org/10.1521/soco.2022.40.4.364>
- Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, 412(6842), 43–48. <https://doi.org/10.1038/35083500>

Open practices statement The data for all experiments are available at <https://doi.org/10.17639/nott.7183>. The materials used for all experiments are available at <https://app.gorilla.sc/openmaterials/451374>. The data analysis files are available at <https://osf.io/zj6h3/>. None of the experiments were preregistered.

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