AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

https://doi.org/10.1037/xge0001190

Instrumental Learning of Social Affiliation Through Outcome and Intention

Hyeji J. Cho and Leor M. Hackel
Department of Psychology, University of Southern California

To build social ties, humans need to find others who want to interact with them. How do people learn, over time, to interact with partners who want to affiliate with them? Theories of social cognition suggest that people try to infer whether others value them, but theories of instrumental learning suggest that rewarding outcomes reinforce choices. In three studies, we provide evidence that both social acceptance outcomes and cues to a partner's acceptance intentions reinforce social partner choices. Even when outcomes were experimentally dissociated from a partner's intentions, outcomes influenced how people felt, which partners people chose, and how well people believed they were liked by partners. Finally, people acted kindlier both to partners who demonstrated acceptance intentions and to partners who provided acceptance outcomes. These findings support an integrative instrumental learning model of social affiliation, wherein social cognition and rewarding outcomes jointly shape affect, partner choice, and prosocial behavior.

Keywords: social cognition, partner choice, social rejection, reward, reinforcement learning

Supplemental materials: https://doi.org/10.1037/xge0001190.supp

Social ties offer a sense of belonging and access to material resources, whereas rejection signals social disregard and bars people from material opportunities. In the short-term, people find social rejection painful (DeWall & Bushman, 2011; Eisenberger et al., 2003; Kross et al., 2011; Leary et al., 1998; Williams et al., 2000), and in the long-term, individuals with more friends are healthier and happier (Cacioppo & Cacioppo, 2014; Holt-Lunstad et al., 2015; Kawachi & Berkman, 2001). As a result, people need to find social partners who will accept them. How do people learn, over time, with whom to affiliate based on experiences of acceptance and rejection?

Leor M. Hackel https://orcid.org/0000-0002-7959-0642

All studies were preregistered, including sample size, measures, exclusion criteria, and analysis plan. Preregistration documents are available at: https://aspredicted.org/qb3zs.pdf, https://aspredicted.org/mb46j.pdf, https://aspredicted.org/zp3vn.pdf. We report all measures, manipulations, and data exclusion criteria. De-identified data and analysis code are available at: https://osf.io/6e73p/?view_only=d3b048ddbf9743f19f72740abe37cfd5. Both authors designed the studies; Hyeji J. Cho collected the data; and both authors contributed to analysis and writing the article. Results from these studies were presented at the Society for Personality and Social Psychology annual meeting in 2021, online, and the Society for Experimental Social Psychology meeting in 2021, Santa Barbara, California. We thank P. Mende-Siedlecki, D. Kalkstein, and members of the USC Social Learning & Choice Laboratory for insightful comments on the article.

Correspondence concerning this article should be addressed to Leor M. Hackel, Department of Psychology, University of Southern California, 3620 South McClintock Avenue, Los Angeles, CA 90089, United States. Email: lhackel@usc.edu

Building a relationship requires repeated actions: People mail an invitation, send a text, or suggest a collaboration. In return, they experience feedback—a friendly response or silence, an RSVP or a rejection—and decide what to do similarly or differently next time. This cycle can be characterized by models of instrumental learning, wherein people perform actions, receive rewarding or punishing feedback, and adjust subsequent behavior (Balleine & Dickinson, 1998; Daw et al., 2011; Sutton & Barto, 2018; Thorndike, 1911). To the extent that people find acceptance rewarding and rejection punishing—given that acceptance fulfills belonging needs and rejection poses a threat to those needs (Baumeister & Leary, 1995; Williams et al., 2000)—people may learn with whom to affiliate through instrumental learning. Yet, rather than serving as a single type of reward, social acceptance includes two distinct types of feedback that could drive learning: acceptance both offers people a concrete outcome of connection with others and reveals that others value them. Here, we dissociate these forms of feedback, integrating approaches from social cognition and reinforcement learning to investigate how people learn which social bonds to build across time.

Acceptance and Rejection as Barometers of Social Value

Social acceptance reveals another person's *intentions* toward us—whether they like us, value us, and prefer to interact with us as opposed to others. Accordingly, people search for cues to their "relational value" in the eyes of others—the degree to which others regard a relationship with them as valuable (Leary, 1999, 2005). In this view, rejection hurts precisely because it reveals that others think poorly of us, which forebodes poor future chances of social connection. Given that humans survive by living in social groups—which requires acceptance from group members—people

vigilantly attend to threats to social connection (DeWall et al., 2009; Leary, 2005). When people perceive that they are not valued, they tend to feel hurt, angry, or lonely (Buckley et al., 2004; DeWall & Bushman, 2011; Leary et al., 1998; Williams et al., 2000). They also adapt their behavior: They tend to seek reconnection with others when possible and avoid or retaliate against those who directly rejected them (Bourgeois & Leary, 2001; DeWall & Richman, 2011; Maner et al., 2007; Twenge et al., 2001).

In contrast, when people do feel valued, they can anticipate that others will accept them over the long-term, fostering closeness and interdependence. Over the course of a relationship, partners are likely to experience conflict or to depend on one another in ways that reveal vulnerability. In these times, people need to know that their partners will not reject them. People therefore seek assurances that they are irreplaceable in a close partner's eyes, and people value partners who are uniquely committed to them or who find them uniquely desirable (Eastwick et al., 2007; Murray et al., 2006). People also feel safer showing vulnerability when they feel assured of a partner's regard for them (Murray et al., 2009; Murray & Holmes, 2009). In this manner, feeling valued encourages people to invest in close relationships, allowing them to feel secure in long-term trajectories of acceptance across relational ups and downs.

Crucially, perceiving one's relational value in another's eyes involves attributing mental states to the other person. People attribute preferences, intentions, and feelings to make sense of others' behavior and predict others' future actions (Heider, 1958). In the case of rejection, people must represent another person's mental states to make sense of the rejection and predict future chances of acceptance. For instance, people might infer that a coworker who rejects a lunch invitation does not care for them—an inference about the coworker's preferences. Given that people generally expect others to act in accordance with preferences (Jara-Ettinger et al., 2016), they might predict that the coworker is unlikely to invite them to a party.

In this manner, people's tendencies to approach or avoid others may depend on mental state attribution following acceptance and rejection. Through interactions over time, people may learn the "relational value" others ascribe to them, and updating this mental state representation may shape who they approach or avoid. People would thus learn to affiliate with individuals who display a desire to interact with them and avoid individuals who do not display this desire—a form of instrumental behavior rooted in modeling others' minds.

Acceptance and Rejection as Positive and Negative Outcomes

Despite the importance of mental state attribution, however, the intentions others bear can diverge from the outcomes others provide. A person might be rejected from joining a team but ranked highly (negative outcome, positive intention) or invited to a party begrudgingly (positive outcome, negative intention). In these cases, acceptance and rejection may reflect not another person's intentions but rather situational constraints (e.g., constraints on the number of team members) or side effects of other goals (e.g., a party planner who wants to invite a guest and feels obligated to invite the guest's significant other). If people respond only to the mental states of others when responding to rejection, then they would ignore these outcomes; one would be ill-served by retaliating against a caring rejecter or by cozying up to a begrudging accepter. Nonetheless, acceptance outcomes may still feel better than rejection outcomes, all else held equal. For instance, a person might feel miffed about being left out of a friend's small wedding, even if their exclusion reflects the event's budget rather than the friend's preferences.

Theories of instrumental learning predict that people repeat actions that yield rewarding outcomes. Such learning has been characterized through models of reinforcement: When people perform an action and receive more rewarding outcomes than they expected, they update a representation of that action's value (i.e., anticipated reward) and become more likely to perform it again (Daw et al., 2011; Sutton & Barto, 2018). Although studies of reinforcement learning in humans have focused primarily on economic reward in nonsocial settings, people also learn to associate other humans with rewarding outcomes. For instance, when people receive economic goods (like monetary gifts) or social goods (like positive feedback) from another person, these outcomes activate neural regions associated with reward (e.g., ventral striatum) and lead people to value their benefactors (Bhanji & Delgado, 2014; Hughes et al., 2018; Jones et al., 2011; Lin et al., 2012; Tamir & Hughes, 2018).

Socioemotional outcomes might similarly reinforce partner choice: If people feel more positive affect after acceptance outcomes, then this experience of psychological reward might lead them to affiliate with individuals who provide bottom-line outcomes of acceptance, regardless of the intentions behind those outcomes. Notably, instrumental learning might lead people to affiliate with others in a manner that diverges from their goals. Instrumental learning can give rise to persistent choice patterns wherein people repeat actions that previously led to reward, even if these actions are no longer relevant to one's goals (Balleine & Dickinson, 1998; Daw et al., 2011; Miller et al., 2019; Wood & Rünger, 2016). Similarly, people may persistently return to partners who previously provided acceptance outcomes, even when those outcomes do not reflect a partner's preferences and would no longer predict a partner's behavior.

Finally, acceptance outcomes may even color how people perceive their relational value in the eyes of others. People often misattribute affect to salient causes (Schwarz & Clore, 1983). If people feel positive affect after acceptance outcomes, then they might associate this affect with a partner and misattribute it to a partner's intentions. People may also infer their own attitudes from their past behavior (Eagly & Chaiken, 1993; Ouellette & Wood, 1998). If people repeatedly interact with partners who provide acceptance outcomes, then they might infer that they have a positive relationship with these partners. Altogether, people may feel positively disposed toward partners who provide acceptance outcomes and believe they are well-liked by those partners. Analogously, people tend to like partners who have provided large monetary rewards in the past, even when these individuals have not shown more generous intentions than others (Hackel et al., 2019; Hackel et al., 2020). People also reciprocate more with these partners, even when there is no strategic benefit to doing so (Hackel & Zaki, 2018). Given that people tend to act kindly toward those who value them and aggress against those who reject them, acceptance outcomes may similarly shape prosocial behavior as well, even when cues to intentions are held constant.

In this manner, people's tendencies to approach or avoid others —as well as their behavior within social interactions—may depend on reward processing following acceptance and rejection. Through interactions over time, people may learn that they tend to experience acceptance outcomes with some people but not others, and this experience of psychological reward may shape who they approach or avoid. As a result, people may learn to affiliate with individuals who offer bottom-line outcomes of acceptance and avoid individuals who offer bottom-line outcomes of rejection—a form of instrumental behavior rooted in reward reinforcement.

A Hybrid Model of Social Instrumental Learning

Altogether, when choosing social partners, people might learn from acceptance outcomes (whether one was accepted or rejected) in addition to acceptance intentions (whether the other person desired interaction). Some evidence hints at this idea: When people anticipate social acceptance feedback, they show increased activation in brain regions linked to both reward processing and social cognition (Powers et al., 2013). Moreover, people do track whether social rejection violates their expectations (Somerville et al., 2006; Sun & Yu, 2014), consistent with computational models of reinforcement learning (Joiner et al., 2017). Yet, it remains unknown what dimensions of feedback underlie these brain activations (i.e., outcomes or intentions) or how these representations relate to partner choice.

Analogously, past work has found that people learn to choose partners who offer materially rewarding outcomes and who display generous character (Hackel et al., 2015, 2020). In this work, participants learned about ostensible earlier participants ("Deciders") who shared points worth money with later participants. Some Deciders shared many points (i.e., rewarding monetary outcomes), whereas others shared a large *proportion* of available points (i.e., had generous character). Participants chose to interact more with both rewarding and generous Deciders, indicating that both types of feedback reinforced choices. While both forms of learning were characterized by prediction error signals in ventral striatum (a hallmark of reward-based learning; Garrison et al., 2013), trait learning (about generosity) was further associated with neural regions previously implicated in social impression updating (Mende-Siedlecki, 2018; Mende-Siedlecki, Baron, et al., 2013; Mende-Siedlecki, Cai, et al., 2013). This evidence suggests that instrumental learning in social interaction involves both rewarding outcomes and social–cognitive cues. More generally, when trying to predict the actions of others, people learn through both reward-based reinforcement and mental state inference (Hampton et al., 2008; Suzuki et al., 2012; Zhu et al., 2012). Yet, it is unknown whether people similarly learn how others value them from acceptance outcomes and mental state cues.

Overview of Studies

Here, we tested a hybrid model of instrumental learning of social affiliation. Drawing on models of reward-based reinforcement and social cognition, we hypothesized that both acceptance outcomes and acceptance intentions reinforce social choices. As a result, people would be more likely to return to social partners who display a desire to interact with them and to partners who offer acceptance as opposed to rejection—much as material reward can directly reinforce choice and promote action repetition (Daw et al., 2011; Hackel et al., 2019). We further hypothesized that learning from outcomes would give rise to patterns of persistent choice, such that people would return to partners who provided acceptance outcomes even when those outcomes were no longer relevant to their goals.

To test these hypotheses, we examined behavior in new experimental designs that dissociated learning from outcome and intention in acceptance and rejection. In Study 1, participants attempted to match with a partner for an economic game and received feedback indicating how much that partner wanted to interact with them (intentions) and whether they actually matched (outcomes). We examined whether both outcomes and intentions influenced partner choice. In Study 2, we examined whether these effects held true when rejection outcomes signaled social disregard but had no economic consequences. Specifically, participants were guaranteed to play a round of the economic game with a random stranger if they did not match with a partner, thus equating financial outcomes across acceptance and rejection. We hypothesized that outcomes would nonetheless reinforce partner choice above and beyond intentions, due to their socioemotional impact. Finally, Study 3 tested whether outcomes and intentions influence not only *whether* people choose to interact with others but also *how* people choose to interact with others. Specifically, we examined consequences for downstream prosocial behavior. We hypothesized not only that participants would choose partners who provided acceptance outcomes and demonstrated acceptance intentions but also that participants would share more money with them.

Study 1

In Study 1, we asked whether people learn from acceptance outcomes and intentions. Participants completed a social game in which they attempted to match with a partner for a trust-based interaction. On each round, participants received feedback indicating the partner's intention to match with them (how the partner ranked the participant relative to others) and the outcome (whether the participant actually matched, due to circumstances outside of the partner's control). By independently manipulating these variables, we tested how each type of feedback influenced social behavior in a trial-by-trial manner.

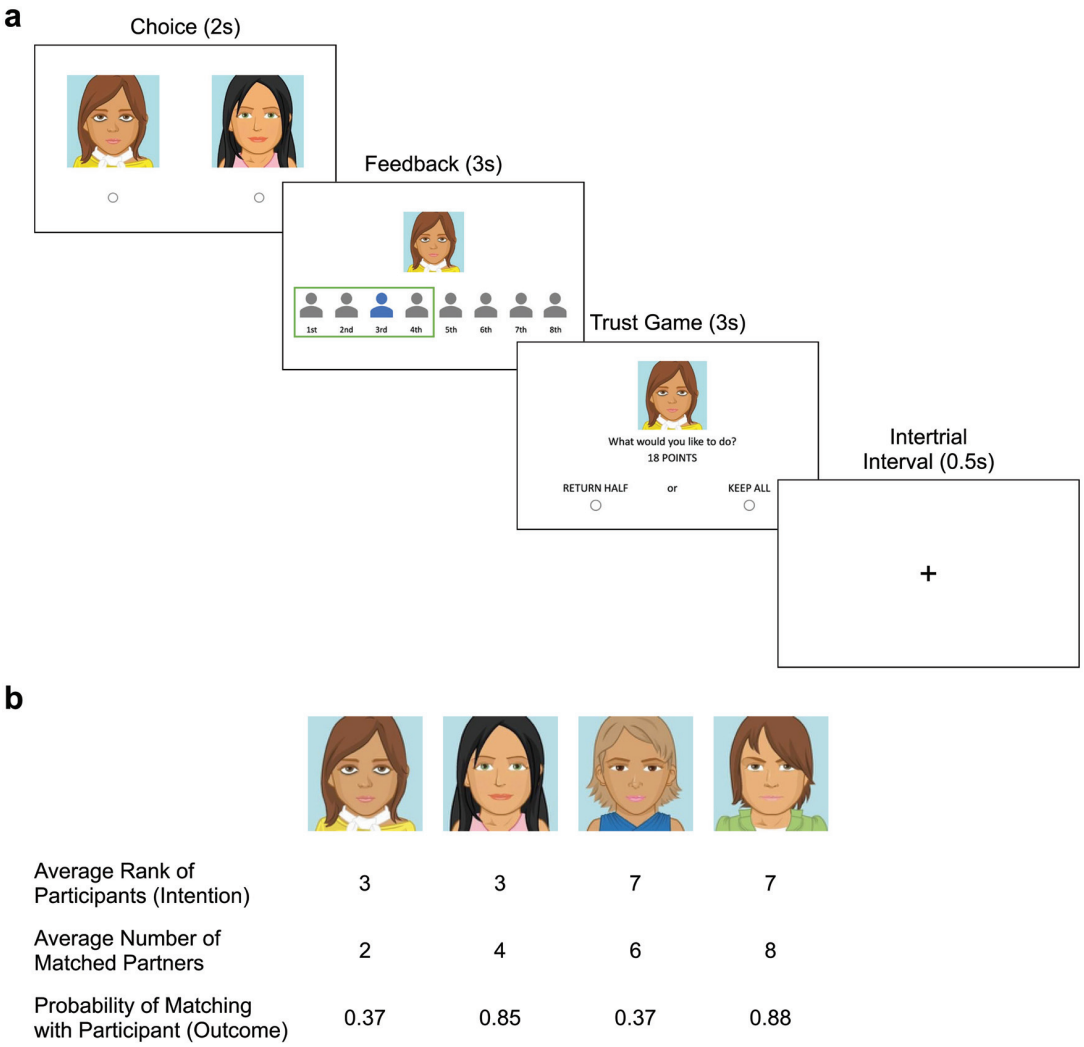
Method

Overview

In an initial session, participants completed a profile about their personalities, including questions relevant to their trustworthiness. A week later, participants were invited back for a second session, in which they learned about previous participants ("Deciders") who had supposedly read their responses, along with the responses of several others, in order to form impressions of each participant. In a subsequent game, participants chose a Decider on each round to try to match with them (Figure 1a). After choosing a Decider, participants received feedback about (a) how that Decider ranked them among eight potential partners (indicating intentions); and (b) whether or not they matched (indicating an outcome). Afterward, participants completed a test phase in which they continued

Figure 1

Schematic of the Learning Task, in Which Participants Learned About Outcomes and Intentions in Social Acceptance and Rejection



Note. (a) On each round, participants first chose a partner and then received feedback about how the partner ranked them (intention) and whether they matched (outcome). The location of the blue avatar indicates the participant's rank (e.g., third), and the size of the green box indicates the number of matches allowed by the computer and thus the participant's outcome (e.g., a matching outcome, given that four matches were allowed). In Studies 1 and 3, participants played a trust-based economic game with that partner only if they matched. In Study 2, participants played with a random other if they didn't match. Face stimuli were used in Study 1 (gender counterbalanced) and animal avatars were used in Studies 2 and 3. (b) Each potential partner varied approximately orthogonally in the average rank (intention) and proportion of matching (outcome) they provided, which was manipulated by varying the average number of matches the computer allotted to each partner on each round (middle row). The images were made at pickaface.net. See the online article for the color version of this figure.

choosing Deciders without feedback under different contingencies that rendered prior acceptance outcomes irrelevant. The test phase thus assessed decisions when previous acceptance outcomes no longer served participants' goals, examining whether participants persisted in choosing Deciders based on prior feedback. Finally, to examine how instrumental learning relates to social perceptions, participants rated the extent to which they believed each Decider liked them.

Participants

In Study 1, 155 participants were recruited on the platform Prolific Academic for Session 1, and 112 of those participants returned for Session 2 (55 women, 55 men, two nonbinary; mean age = 33.53, range = 18 to 77). Prolific has been found to yield high-quality data (Palan & Schitter, 2018). Only participants who completed both Session 1 and Session 2 were included in the analyses. To

ensure that participants were actively engaged in the task, we administered a preregistered exclusion rule to remove data from any participant who did not respond to at least 80% of trials in the learning phase and the test phase. Using these criteria, 15 participants were excluded, leaving 97 participants for analyses. Sample size was chosen to provide at least 80% power to examine correlations with individual difference variables, which were expected to be the smallest effect of interest, at a moderate effect size (r= .30), plus additional subjects to account for exclusions. Informed consent was obtained from all participants in accordance with approval from the USC Office for Protection of Human Subjects.

Stimuli

Decider participants were represented by face avatars (created on the website pickaface.net), which were all male or all female (randomized across participants, in order to allow generalization across gender between subjects without varying gender within subjects). Avatars were randomly assigned to different task roles shown in Figure 1b across participants to minimize any effects of a particular avatar on the results.

Procedure

In Session 1, participants were told that they would be playing a "Getting-to-Know-You" game with other participants. They answered six questions about their personality, particularly focused on their trustworthiness (e.g., "When was a time when you were honest, even though you didn't have to be?"), which they were told would then be sent to other participants to read. Afterward, participants completed questionnaire measures in order to assess individual differences that may be relevant to responses to social rejection (see online supplemental materials).

A week later, participants were invited back for Session 2. In this session, they were told that four other participants in a "Player A" role (termed "Deciders" here) read their responses and the responses of 12 other participants in a "Player B" role (termed "Responders" here). Then, Deciders picked partners from this group of Responders for a game. For every round of the game, a Decider could choose to send points worth money to each of a varying number of Responders. Points would be tripled, and Responders would then choose whether to return half of the points to that Decider or keep all of them. (In reality, all participants were assigned to the Responder role.)

To play the game on a particular round, participants needed to match with a Decider. Participants were informed that each Decider saw a random combination of eight out of 12 Responders on each round and ranked who they wanted to play with from 1 (*most desired*) to 8 (*least desired*); this instruction offered an explanation of why rankings could vary from trial to trial, given different sets of Responders available. Participants were told that a computer next allocated a number of Responders that each Decider could match with on that round, from one to eight matches. For instance, if a Decider was allowed four matches, they would play with their top four ranked Responders; if allowed six matches, they would play with their top six ranked Responders. Deciders could thus have positive intentions toward a participant (e.g., ranking the participant third) but yield negative outcomes (failing to match, if allowed only two partners), or Deciders could have negative intentions (e.g., ranking the participant sixth) but yield positive outcomes (matching, if allowed seven partners). This design is analogous to being chosen for a large team as a last choice versus failing to make a small team but knowing one would have been the next choice. Unbeknownst to participants, by allocating larger sets of matches to some Deciders than others, we manipulated the average rank and average probability of matching provided by each Decider (see Figure 1b for details).

Learning Phase

Across 96 trials, participants learned about the four Decider participants. On each trial, participants had 2 s to choose one of two Deciders (out of the four) shown onscreen by pressing either "E" (left) or "I" (right) on their keyboard. The combination and the location (i.e., left vs. right) of the Deciders displayed on each trial were counterbalanced, and each possible pair appeared an equal number of times. The time window was limited to 2 s in order to standardize the amount of time each participant had available to make decisions while allowing for both goal-directed ("model-based") and less goal-directed ("model-free") forms of choice, as indicated by past work using similar time windows (Doll et al., 2015; Kool et al., 2017). If participants did not make a decision within 2 s, they saw a screen that said "No response" for an additional 3 s before moving on to the next trial. This timing equated the length of trials when the participant did not respond or when they responded but did not match.

After choosing a Decider, participants received 3 s of feedback about how that Decider ranked them relative to seven other Responders (indicating intentions) and whether they matched, based on the number of matches the Decider was allowed (indicating outcomes). A green box showed the number of matches the computer had allowed for that Decider; the green box could include or exclude the participant (a blue avatar to be distinguished from the other Responders with gray avatars), indicating the outcome (Figure 1a). The placement of the participant's avatar in the row of potential Responders indicated their ranking (e.g., first through eighth going from left to right). To generate feedback, Gaussian noise was added to the target's average ranking and average number of matches (SD = 1 for each); a censored distribution was used, such that both quantities had to be between one and eight. If the participant's ranking was within the number of matches allowed, they matched with that Decider for the trial. This procedure produced the average probabilities of matching shown in Figure 1b (as observed in 10,000 simulations of the procedure).

If they matched with the Decider they chose, participants could play the trust game. They were shown a number of points worth money sent from that Decider that had been tripled (ranging from a total of 6 to 24 points). They had 3 s to decide whether to keep all of those points or return half of them to that Decider, indicated by a button press. If they didn't match, they couldn't play the game and instead had to wait 3 s before moving on to the next trial. These trust games during the learning phase served only to maintain the cover story, giving participants a reason to choose partners and offering a consequence for not being chosen in return.

To directly measure the impact of outcome and intention feedback on affect, participants rated their affect in nine trials randomly distributed throughout the task. Specifically, participants were asked to rate how they felt about that round on a scale of 1

(*very bad*) to 7 (*very good*) after receiving feedback but before the trust game.

Test Phase

Participants next completed a test phase of 48 trials in which they continued choosing Deciders without feedback. Participants were told that they would later receive feedback and play trust games based on these choices at the very end of the task. They had 3 s to choose between two Deciders. Similar to the learning phase, pairing and location (i.e., left vs. right) of Deciders shown on each trial were counterbalanced.

However, the test phase introduced a change in the contingencies by which matches were assigned. Specifically, the number of matches allowed to each Decider was explicitly displayed onscreen during decision-making, and Deciders varied uniformly in whether they had many or few matches available on each trial. On half the trials, both Deciders had five matches available, while on the other half of trials, one Decider had three matches available and the other had seven. (Each Decider appeared at each ratio an equivalent number of times in each pairing.) All Deciders therefore had an equivalent number of matches available on average, meaning that rates of acceptance outcomes from the learning phase were not carried over to the test phase.

These changes render earlier outcomes from the learning phase irrelevant to participants' goals. During learning, some Deciders systematically had more matches available than others and therefore provided more matching outcomes regardless of their intentions (Figure 1b). For example, one Decider often ranked participants seventh—a poor ranking—but typically had eight matches available and thus provided matching outcomes. As a result, it would have been reasonable for a participant to choose based on both outcome and intention during learning. In the test phase, however, the number of matches available was known and equated on average across Deciders. As a result, only rankings should be relevant to choice; if two Deciders can each match with four partners, then the Decider who ranks a participant more favorably would be more likely to offer a match. This structure allowed us to test whether participants were sensitive to the new contingencies, as indicated by choosing Deciders who had more matches displayed onscreen, while also testing whether participants showed persistent choice patterns, as indicated by continuing to choose Deciders who provided more positive outcomes during learning.

Liking Ratings

Finally, to measure participants' explicit perceptions of the Deciders, participants completed two sets of ratings at the very end. On a scale from 1 (*not at all*) to 7 (*very much*), participants rated the extent to which they believed each Decider liked them and rated the extent to which they liked each Decider (see Table S6). The order in which they saw each Decider and the order in which they saw each set of ratings were randomized.

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All studies were preregistered, including sample size, measures, exclusion criteria, and analysis plan. Preregistration documents are available at: https://aspredicted.org/qb3zs.pdf, https://aspredicted.org/mb46j.pdf, https://aspredicted.org/zp3vn.pdf. De-identified data and analysis code are available at: https://osf.io/6e73p/?view_only=d3b048ddbf9743f19f72740abe37cfd5 (Cho, 2021).

Results

Learning

We first asked whether participants learned from outcome feedback (attaining a match), intention feedback (being ranked highly), or both. We fit behavior to a computational model adapted from models of reinforcement learning in social interactions (Hackel et al., 2015, 2020). Conceptually this model updates estimates of expected rankings and expected outcomes using prediction errors for each (i.e., the difference between feedback received and feedback expected). To make decisions, these estimates are combined as a weighted average using a weighting parameter *w* ranging continuously between 0 (*fully outcome-based*) and 1 (*fully intention-based*).

More specifically, this model assumes that participants update a reward value Q and intention estimate I following feedback on each trial t according to:

$$Q_t = Q_{t-1} + a_R \delta_{Rt} \tag{1}$$

$$I_t = I_{t-1} + \alpha_I \delta_{It} \tag{2}$$

where α_R and α_I are free parameters representing learning rates for outcome and intention, respectively; δ_{Rt} represents a reward prediction error on trial t; and δ_{It} represents an intention prediction error on trial t. Prediction errors are defined as the difference between values received and values expected for reward and intentions:

$$\delta_{Rt} = Reward - Q_{t-1} \tag{3}$$

$$\delta_{It} = Intention - I_{t-1} \tag{4}$$

Reward was a binary value indicating acceptance (1) or rejection (0), and intention was defined as the ranking given. An expected value based on intention was defined as the probability of matching given expected rankings, assuming a uniform probability of group sizes from 1 to 8:

$$IV = 1 - \frac{I - 1}{8} \tag{5}$$

In this manner, estimates of intentions were agnostic as to average outcomes and estimates of outcome were agnostic to average intentions. Outcome expectations were initialized to .50 and intention expectations were initialized to the midpoint of the possible rankings (4.5) to represent initial uncertainty.

The model allowed integration of intention-based values and reward-based values into an overall expected value according to:

$$EV = w(IV) + (1 - w)Q \tag{6}$$

where w is a weighting parameter indicating how much participants rely on intention values or outcome values. A participant who relies only on intentions would have a weighting parameter

w = 1, while a participant who relies only on outcomes would have a weighting parameter w = 0.

Finally, participant choices were modeled using a softmax choice function:

$$p_{i,t} = \frac{\exp(\beta \times EV_{i,t})}{\sum_j \exp(\beta \times EV_{j,t})} \quad (7)$$

where β is an exploration parameter controlling stochasticity of choice in the learning phase and $p_{i,t}$ is the probability of choosing option i (of j options) on trial t.

This model thus had four free parameters: α_R, α_I, w, and β (see Table S2 for parameter fits). Parameters were estimated using maximum a posteriori (MAP) estimation to optimize parameters across all choices, using priors of gamma(1.2, scale = 5) applied to exploration parameters and beta(1.1, 1.1) applied to learning rates and the weighting parameter (Daw et al., 2011; Decker et al., 2015; Hackel et al., 2015, 2020).

For each model, the best-fitting parameters were used to compute the Laplace approximation to the Bayesian model evidence. Two alternative models were tested: An intention-only model, in which the w parameter was fixed to 1 and α_R was fixed to 0, and an outcome-only model, in which the w parameter was fixed to 0 and α_I was fixed to 0. These models had two fewer free parameters than the hybrid model. Models were compared using random effects Bayesian model comparison (Stephan et al., 2009), implemented via the spm_bms function in the SPM12 toolbox for Matlab, as well as fixed effects methods (see Table S1). To validate the model, we further tested its ability to reproduce qualitative patterns of behavior in the learning phase and its ability to predict out-of-sample choices in the test phase (supplemental methods, Figures S1–S2).

Computational modeling of behavior revealed that participants learned from both outcomes and intentions. Random effects Bayesian model comparison favored the hybrid model over simpler models that included only outcomes or intentions (exceedance probability = 1; see Table S1 for fixed effects Bayesian model comparison and Table S2 for parameter estimates). The median w parameter similarly indicated reliance on both outcome and intention (median = .49). Exploratory analyses further revealed that learning rates for intentions (median = .44) were significantly higher than learning rates for outcomes (median = .21), z = –2.93, $p = .003$ (sign rank test due to non-normality), further dissociating these forms of learning. That is, upon receiving feedback, participants updated estimates of others' intentions more strongly than they updated estimates of likely outcomes.

Results of the computational model were further supported by a mixed effects regression analysis, which approximated the full reinforcement learning model through traditional linear analysis methods (see supplemental methods). This analysis revealed main effects of both outcome and intention (but no significant interaction between them; see Table S3). Thus, participants learned to interact with Deciders who ranked them highly and Deciders with whom they successfully matched.

Affect

This finding coheres with the proposal that positive outcomes and positive intentions each influence how people feel after acceptance feedback, even when outcomes are dissociated from a partner's intentions. To directly test whether outcomes and intentions influenced how participants felt, we analyzed affect ratings during the learning phase using mixed-effects linear regression, in light of repeated measures within subjects. The model predicted affect ratings as a function of that trial's outcome (1 = match, –1 = no match) and ranking (continuous, reverse scored such that higher values indicate more positive intentions and standardized to z-scores within subjects). Fixed and random effects were allowed for all predictors. Analyses were performed using the lme4 and lmerTest packages for R (Bates et al., 2014; Kuznetsova et al., 2017). As a measure of effect size for linear mixed effect regression models, we report semipartial R^2 for fixed effects predictors (Edwards et al., 2008) as computed using the r2glmm package for R (Jaeger, 2017).

Indeed, participants reported more positive affect after good versus bad rankings, $b = .22$, $SE = .06$, $t(115.90) = 3.56$, $p < .001$, 95% CI [.10, .33], $R^2_\beta = .10$, but also after matching versus nonmatching outcomes, $b = .40$, $SE = .07$, $t(97.02) = 5.74$, $p < .001$, 95% CI [.26, .54], $R^2_\beta = .25$. Thus, both cues to intentions and bottom-line outcomes of acceptance led participants to feel positive affect, generating experiences of psychological reward. An interaction between outcome and intention suggested that intentions particularly impacted affect during positive outcomes, $b = .11$, $SE = .05$, $t(741.94) = 2.37$, $p = .02$, 95% CI [.02, .21], $R^2_\beta = .01$ (although this interaction did not replicate in Studies 2–3; see Table S4).

Test Phase

In light of this impact on affect and choice, instrumental learning from social outcomes might lead people to persistently choose partners who previously accepted them, even if doing so is no longer necessary for one's goals, as can be the case with materially rewarding outcomes (Balleine & Dickinson, 1998; Daw, 2011; Hackel et al., 2019; Wood & Rünger, 2016). To test this hypothesis, test phase choices were analyzed using mixed effects logistic regression, in order to model trial-by-trial choices. The model predicted the probability of choosing the Decider on the right side of the screen (arbitrarily chosen) as a function of difference in average rank provided by each Decider (right minus left average rank), difference in average outcomes provided by each Decider (right minus left average probability of matching), and the difference in number of matches available onscreen for each Decider (right minus left number of matches). In order to analyze data in a manner independent of the computational model, targets were scored as high (1) or low (–1) on each variable.

Participants were sensitive to the new contingencies, choosing Deciders who had more matches available during test phase trials, $b = .40$, $SE = .05$, z = 7.43, $p < .001$, $OR = 1.49$, 95% CI [1.34, 1.66], and choosing Deciders who had ranked them highly during learning, $b = 1.32$, $SE = .18$, z = 7.20, $p < .001$, $OR = 3.74$, 95% CI [2.61, 5.35]. Thus, participants were aware of the new contingencies and, consistent with those continencies, chose Deciders who had ranked them highly and who now had many matches available. Strikingly, however, participants also continued to choose partners who had previously provided more frequent matching outcomes during the learning phase, $b = 1.27$, $SE = .15$, z = 8.71, $p < .001$, $OR = 3.57$, 95% CI [2.68, 4.75], suggesting a persistent impact of acceptance outcomes on decision-making

(Figure 2a). No significant interaction was observed between outcome and intention, $b = .06$, $SE = .07$, $z = .81$, $p = .418$, $OR = 1.06$, 95% CI [.92, 1.23].

Perceived Liking

We next asked whether intentions and outcomes influenced participants' perceptions of how well they were liked. Ratings of perceived liking from each Decider were entered into a 2 (rank: high, low) × 2 (outcome: high, low) repeated measures ANOVA. Indeed, although participants strongly perceived they were better liked by Deciders who ranked them highly, $F(1, 96) = 64.93$, $p < .001$, $\eta_p^2 = .40$, they also perceived that they were better liked by Deciders who matched with them more often, $F(1, 96) = 36.73$, $p < .001$, $\eta_p^2 = .28$ (Figure 3a; Table S7). No interaction between outcome and intention was observed, $F(1, 96) = .20$, $p = .66$, $\eta_p^2 = .002$. Acceptance outcomes thus shaped participant's perceptions of being liked, even when two Deciders were explicitly shown to have equal preferences toward the participant.

To better understand the contribution of instrumental learning to social perception, we examined the extent to which perceptions of being liked reflected outcomes and intentions. Specifically, we computed an individual difference measure of sensitivity to outcomes versus intentions in social perceptions. To do so, we first computed difference scores representing "intention-based perceptions," collapsing across outcome levels:

Intention-based perception = [average ratings for positive ranking Deciders] – [average ratings for negative ranking Deciders]

Similarly, we computed difference scores indicating "outcome-based perceptions," collapsing across intention levels:

Outcome-based perception = [average ratings for frequently matching Deciders] – [average ratings for infrequently matching Deciders]

Finally, we computed the relative strength of intention versus outcome by taking a difference score of these two measures:

Perception difference score = [intention-based perception – outcome-based perception]

As is true of the w parameter, higher scores indicate relatively greater reliance on intentions and lower scores indicate relatively greater reliance on past outcomes.

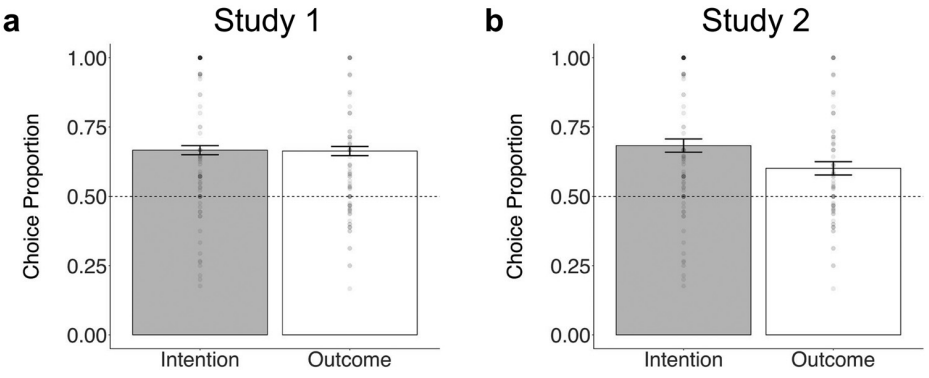
In order to test whether participants relied more on intentions than outcomes on average in these judgments, a one-sample t-test was used to compare the difference score with zero (Table S5). Notably, the effect of intentions was stronger than the effect of outcomes, $t(96) = 2.41$, $p = .02$, $d = .25$, indicating that participants primarily inferred liking based on rankings, which provided a meaningful signal to the Decider's intentions, and secondarily formed perceptions based on outcomes, which were independent of the Decider's intentions.

However, the extent to which participants relied on each form of feedback when assessing how well they were liked varied with their learning styles. Specifically, we examined the correlation between the weighting parameter (w) extracted from the computational model and difference scores for perceptions of liking. Individuals who relied more on outcome feedback (vs. intention feedback) when choosing partners during the learning phase, as indexed by the model's weighting parameter w, also relied more on outcomes (vs. intentions) when judging how much they were liked, $r(95) = .43$, $p < .001$, 95% CI [.25, .58] (Figure 3d). Thus, individual differences in instrumental learning explained variance in perceived liking, directly linking learning and social perception.

Discussion

In Study 1, we investigated how people learn to interact with partners based on social acceptance feedback. We found that both outcome and intention feedback influenced affect and reinforced partner choices: Participants felt better and were more likely to return to partners when they were ranked highly and when they

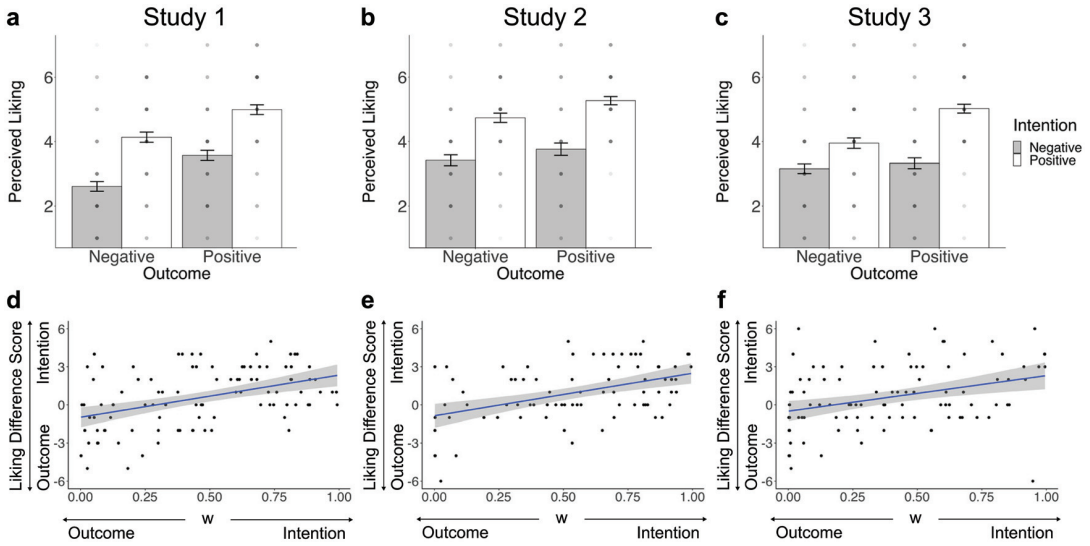
Figure 2
Test Phase Choices Based on Intention and Outcome



Note. Plots show the proportion of choices for which participants selected the partner onscreen that previously gave higher rankings (intention), and, independently, the proportion of choices for which participants selected the partner that previously offered more frequent matches (outcome). Dots indicate individual data points, with darker shade indicating a higher density of points. Error bars indicate standard error of the mean, with within-participant adjustment (Morey, 2008). (a) In Study 1, both intention and outcome predicted partner choice above chance (dotted line) even when contingencies changed, such that participants continued to choose partners who ranked them highly and who had previously provided more frequent matching outcomes. (b) Study 2 replicates the findings from Study 1 when outcomes carried only socioemotional, rather than material value.

Figure 3

Explicit Ratings of Participants' Perceptions of How Well They Were Liked by Each Partner



Note. Dots indicate individual data points, with darker shade indicating a higher density of points. Error bars indicate standard error of the mean, with within-participant adjustment (Morey, 2008). (a) Participants perceived that they were better liked by partners associated with more positive intentions and more positive outcomes in Study 1, (b) Study 2, and (c) Study 3. (d) Participants who relied more on outcomes (versus intentions) when choosing partners during the learning phase, as revealed by the model's weighting parameter (w), also relied more on outcomes (versus intentions) when judging how much they were liked in Study 1, (e) Study 2, and (f) Study 3. See the online article for the color version of this figure.

received a matching outcome. Participants thus updated representations of their partners' mental states (i.e., intention; how much the partner desired interaction with them) and of the direct reward value of choosing a particular partner (i.e., outcome; the likelihood of acceptance). They used both representations to make choices.

Notably, the effect of outcome persisted when previous outcomes were rendered irrelevant: participants continued to choose partners who offered positive outcomes even when those prior outcomes were no longer relevant to decisions. Moreover, participants thought they were better liked not only by those who ranked them highly but also by those who provided more positive outcomes, even though partners only controlled rankings. These findings are consistent with the idea that acceptance outcomes serve as a psychological reward that promotes action repetition, as in other forms of instrumental learning, and that this experience of reward can lead people to view others more positively. Together, these findings support an instrumental learning model of social affiliation that incorporates both outcome and intention and begin to link instrumental learning to social perception.

Study 2

In Study 2, we aimed to replicate Study 1 while removing any economic incentives of matching with partners. Although Study 1 dissociated learning from acceptance outcomes and intentions, it remained possible that the effects of outcomes depended on economic incentives, as in typical studies of reinforcement learning. Specifically, in Study 1, participants who failed to match with a Decider did not get to play the trust game on that round, thus losing an economic opportunity. In Study 2, participants were therefore allowed to play the trust game with a random other person on unmatched trials. This change eliminated any economic incentive to match, allowing a test of whether socioemotional outcomes alone are sufficient to influence learning.

Method

Participants

In Study 2, 125 participants were recruited on Prolific for Session 1, and 93 of those participants (49 women, 43 men; mean age = 31.06, range = 18 to 65) returned for Session 2. Using the same exclusion criteria as Study 1, an additional fourteen participants were excluded, leaving 79 participants for analyses. Informed consent was obtained from all participants in accordance with approval from the USC Office for Protection of Human Subjects.

Procedure

The procedure of Study 2 was identical to that of Study 1, with three exceptions. First, Decider participants were represented by colorful animal avatars, similar to those used on prominent collaboration websites; these avatars were used to minimize cues to social categories (e.g., gender) that might influence learning. As in Study 1, avatars were randomly assigned to Decider roles shown in Figure 1b across participants to minimize any effects of a particular avatar on the results.

Second, in the learning phase of Study 2, participants were told that if they didn't match with the Decider they chose on a given round, they could still play the trust game with a random other participant who also didn't get matched. They thus played a trust

game on every round, regardless of outcome feedback, meaning that economic incentives remained consistent across matching and nonmatching outcomes.

Finally, in the test phase of Study 2, all Deciders had four matches available on each round. The purpose of this change was to simplify the test phase; having already demonstrated sensitivity to the number of matches available in Study 1, we aimed to simplify the processing demands of the task by making clear that all Deciders always had an equivalent number of matches. Again, if all Deciders always have an equivalent number of matches available, then the optimal strategy would be to choose Deciders solely based on rankings they provide.

Results

Learning

We fit choices during the learning phase to the same computational models described in Study 1 and performed model comparison to determine whether partner choice was best predicted by intentions, outcomes, or both. Results of this study replicated those of Study 1. Specifically, participant choices were best fit by a reinforcement learning model that included not only intentions but also outcomes (exceedance probability = 1; see Tables S1–S2). The weighting parameter w again indicated learning from both outcome and intention (median = .55). Again, learning rates for intentions (median = .46) were significantly higher than learning rates for outcomes (median = .27), z = –2.50, $p = .01$ (sign rank test), indicating faster learning from intentions than outcomes. Finally, mixed effects regression analyses again supported the conclusion that participants learned from outcomes and intentions (with no significant interaction between the two forms of feedback; Table S3). These results demonstrate that socioemotional outcomes were sufficient to reinforce choices in the absence of economic incentives.

Affect

To understand the impact of outcome and intention feedback on participant feelings, we again used a mixed effects linear regression to predict affect ratings as a function of each type of feedback. Participants once again felt more positive affect after receiving acceptance than rejection outcomes, $b = .18$, $SE = .05$, $t(84.92) = 3.61$, $p < .001$, 95% CI [.08, .28], $R^2_\beta = .13$, in addition to feeling more positive after receiving high as opposed to low rankings, $b = .25$, $SE = .06$, $t(92.93) = 4.04$, $p < .001$, 95% CI [.13, .37], $R^2_\beta = .15$ (Table S4). Thus, participants felt more positive affect when they were accepted—a socioemotional outcome—even when intentions and monetary reward were held constant. Unlike Study 1, no significant interaction was observed between outcomes and intentions, $b = .02$, $SE = .05$, $t(618.97) = .37$, $p = .715$, 95% CI [–.08, .12], $R^2_\beta = < .001$.

Test Phase

We next asked whether acceptance outcomes were sufficient to prompt persistent choice patterns in the test phase, even when outcomes had no financial consequences for participants. Indeed, using the same mixed-effects logistic regression analysis as described in Study 1, we found that participants persistently chose Deciders who provided acceptance outcomes in a subsequent test phase with new contingencies, $b = .67$, $SE = .15$, z = 4.44, $p < .001$, $OR = 1.96$, 95% CI [1.46, 2.63], in addition to choosing Deciders who provided high rankings, $b = 1.40$, $SE = .24$, z = 5.92, $p < .001$, $OR = 4.06$, 95% CI [2.55, 6.45] (Figure 2b). Again, we found no interaction between outcomes and intentions in test phase choice, $b = .10$, $SE = .08$, z = 1.24, $p = .217$, $OR = 1.10$, 95% CI [.94, 1.28]. Thus, participants continued choosing Deciders who had offered acceptance outcomes, even when these outcomes were no longer relevant to their goals and when outcomes carried only socioemotional weight.

Perceived Liking

We again examined participants' ratings of how well they were liked by each Decider. Ratings were again submitted to a 2 (Rank: High, Low) × 2 (Outcome: High, Low) repeated measures ANOVA. Replicating Study 1, participants inferred they were liked not only by Deciders who had ranked them highly, $F(1, 78) = 49.92$, $p < .001$, $\eta^2_p = .39$, but also by Deciders associated with frequent acceptance outcomes, $F(1, 78) = 10.67$, $p = .002$, $\eta^2_p = .12$ (Figure 3b, 3e). Again, no interaction was observed between outcome and intention, $F(1, 78) = .60$, $p = .44$, $\eta^2_p = .01$.

Finally, we again tested the extent to which participants' perceptions of being liked relied on ranking feedback, which provided a meaningful cue to the Deciders' mental states, and outcome feedback, which did not. We computed the same difference score described in Study 1 reflecting relative reliance on each form of feedback, with higher scores representing reliance on intentions and lower scores representing reliance on outcomes. The mean difference score was again significantly greater than zero, $M = .95$, $t(78) = 3.85$, $p < .001$; d = .43, indicating that participants primarily relied on rankings as a cue to being liked and secondarily relied on outcomes. Moreover, the w parameter during learning again predicted subsequent reliance on intentions, relative to outcomes, in perceptions of being liked, $r(77) = .45$, $p < .001$, 95% CI [.26, .61] (Figure 3e). Participants who relied more on outcomes (vs. intentions) when choosing partners during the learning phase also relied more on outcomes (vs. intentions) when judging how much they were liked, once again linking differences in learning to differences social perception.

Discussion

In Study 2, we replicated the results of Study 1 while using strictly socioemotional—as opposed to economic—reward. Participants played an adapted game in which rejection had no material consequences. Despite this change, both outcome and intention feedback continued to reinforce partner choice, shape participant affect, and produce perceptions of being liked. Even though outcomes carried only socioemotional weight, outcomes still gave rise to persistent patterns of choice: Participants continued choosing partners who previously provided matching outcomes even when these outcomes were no longer goal-relevant. These findings expand models of instrumental learning in social interaction, demonstrating that purely socioemotional feedback from social acceptance or rejection shapes instrumental learning and choice.

Study 3

In Study 3, we asked whether acceptance outcomes and intentions influence not only whether people choose to interact with others but also how people choose to interact with others. Studies 1 and 2 examined the extent to which people approach or avoid others in response to intention and outcome feedback. However, people also treat others differently following acceptance or rejection: People often act kindly toward those who accept them and retaliate against those who reject them (DeWall & Bushman, 2011; Twenge et al., 2001). (Although people do sometimes act prosocially after rejection, they tend to do so only when they believe they can regain connection, and they tend to direct these efforts toward others besides the individuals who directly rejected them; Maner et al., 2007). Prosocial and retaliatory responses could depend on intentions, outcomes, or both. If acceptance outcomes make people feel good and believe they are liked, then people may act kindlier to those who accept them than to those who reject them, even when intentions are held constant. This tendency could produce unwarranted social conflict when rejection outcomes do not reflect rejection intentions.

To test this possibility, Study 3 featured a prosocial choice phase after the learning phase. Participants completed additional rounds of a continuous trust game with each of the four Deciders they had learned about. On each round, they saw one of the Deciders along with money sent from that Decider; as in the learning phase, Deciders always sent their full endowment. Participants then indicated how much to repay on a continuous scale. This phase therefore examined prosocial choices due to outcome and intention feedback after learning was complete and when participants had an equal number of opportunities to interact with each Decider.

Method

Participants

In Study 3, 125 participants were recruited on Prolific for Session 1, and 99 of those participants (51 women, 44 men, four nonbinary; mean age = 29.98, range = 18 to 67) returned for Session 2. The same exclusion rule was applied as in previous studies, but using only learning phase data, given that there was no test phase. Using this rule, an additional four participants were excluded, leaving 95 participants for analyses. Informed consent was obtained from all participants in accordance with approval from the USC Office for Protection of Human Subjects.

Procedure

The procedure of Study 3 was identical to that of Study 1, with two exceptions. First, animal avatar stimuli were used as in Study 2. Second, participants completed an additional set of trust games instead of the test phase, allowing us to cleanly measure prosocial behavior on the basis of prior feedback. Although participants did make trust decisions during the learning phases of Studies 1 and 2, these decisions did not permit a clear test of intention and outcome for several reasons. First, participants were still learning about Deciders during the learning phase. Second, participants had fewer chances to make trust decisions with low-outcome Deciders who provided few matches, by definition. Third, trust choices in the learning phase were binary, offering low resolution to detect differences across Deciders. Therefore, in Study 3, participants completed a dedicated prosocial choice phase after learning was complete, in which they saw each Decider an equal amount of times and made continuous decisions to repay any amount between zero and the full endowment.

Participants were told that they would again be playing with each of the four Deciders, but that Deciders were now allowed to play the game with all eight Responders available on every round. Participants therefore did not choose who to play with or receive any feedback during this stage. Instead, on each round, they saw a point total sent from a Decider and chose how much to return using a slider scale ranging from zero to the maximum amount. Participants played 20 trials in total, with each of the four Deciders appearing five times with different point totals in each repetition (offering the possibility of returning up to 30, 45, 60, 75, or 90 points).

Results

Learning and Affect

The learning phase of Study 3 again replicated all previous findings across choice and affect (Tables S1–S5). Participant choices were best fit by the hybrid model of learning (exceedance probability = 1), indicating that they learned from outcomes and intentions, and this inference was supported by a mixed effects regression analysis revealing main effects of each type of feedback (with no significant interaction between them; Table S3). Again, the learning rate for intentions (median = .47) was significantly higher than the learning rate for outcomes (median = .22), $z = -4.06$, $p < .001$ (sign rank test). Similarly, in a mixed effects regression analysis of affect ratings, participant affect depended on both intention feedback, $b = .15$, $SE = .06$, $t(118.25) = 2.60$, $p = .01$, 95% CI [.04, .27], $R^2_\beta = .05$, and outcome feedback, $b = .38$, $SE = .05$, $t(94.44) = 8.10$, $p < .001$, 95% CI [.29, .48], $R^2_\beta = .41$, with no significant interaction observed between them, $b = -.04$, $SE = .04$, $t(778.50) = -.87$, $p = .38$, 95% CI [–.13, .05], $R^2_\beta = .001$.

Perceived Liking

Analyses of liking ratings similarly replicated findings of Studies 1 and 2 (Figure 3c). Participants perceived they were better liked by Deciders who ranked them highly, $F(1, 94) = 42.73$, $p < .001$, $\eta^2_p = .31$, and who matched with them more often, $F(1, 94) = 19.53$, $p < .001$, $\eta^2_p = .17$. Unlike previous studies, an Intention × Outcome interaction suggested that rank had a larger impact for Deciders who frequently offered positive outcomes, $F(1, 94) = 19.99$, $p < .001$, $\eta^2_p = .18$. As in prior studies, the effect of intention was greater than the effect of outcome, $t(94) = 2.39$, $p = .02$, $d = .25$. Additionally, participants who relied more on outcomes (vs. intentions) when choosing partners during the learning phase, as revealed in the model's weighting parameter (w), also relied more on outcomes (vs. intentions) when judging how much they were liked, $r(93) = .34$, $p < .001$, 95% CI [.15, .51] (Figure 3f).

Prosocial Choice

To examine whether outcomes and intentions influence not only whether people interact with others but also how people interact with others, we examined participants' choices during the prosocial

choice phase. Participant choices were fit to a mixed effects linear regression model. This model predicted the proportion of points participants returned on each round as a function of the Decider's average outcome (1 = high, –1 = low) and average rank (1 = high, –1 = low) during learning. Proportion was used, rather than absolute amount, because participants saw different point totals on different trials. As in other analyses, fixed and random effects were allowed for all predictors.

When making prosocial choices, participants indeed returned more money not only to Deciders who gave them high rankings, $b = .04$, $SE = .01$, $t(94.00) = 3.62$, $p < .001$, 95% CI [.02, .05], $R^2_\beta = .12$, but also to Deciders associated with more frequent matching outcomes, $b = .02$, $SE = .01$, $t(93.95) = 2.74$, $p = .007$, 95% CI [.004, .03], $R^2_\beta = .07$, even though Deciders only controlled intentions (Figure 4a). No significant interaction between outcomes and intentions was observed, $b = .01$, $SE = .004$, $t(94.04) = 1.64$, $p = .10$, 95% CI [–.001, .02], $R^2_\beta = .03$. These findings demonstrate that instrumental learning from outcome and intention gives rise not only to patterns of approach versus avoidance but also to patterns of prosocial versus retaliatory behavior.

To further understand the extent to which each form of learning gave rise to prosocial behavior, we computed a difference score indicating the extent to which participants shared with Deciders based on rankings versus outcomes; this difference score was computed for proportions shared using a procedure identical to that used with ratings of being liked. We examined the correlation between this difference score and the w parameter extracted from the learning model. Indeed, the w parameter during learning predicted subsequent reliance on intentions, relative to outcomes, in prosocial behavior, $r(93) = .28$, $p = .006$, 95% CI [.08, .46] (Figure 4b)[1]: Individuals who relied more on outcome feedback (vs. intention feedback) when choosing partners during the learning phase also relied more on outcomes (vs. intentions) when repaying money in the trust game, further linking instrumental learning along these two dimensions to prosocial behavior.

Discussion

Beyond replicating the findings of Studies 1–2, Study 3 investigated whether acceptance outcomes and intentions influence prosocial behavior. When presented with an opportunity to send money to others, participants sent more money to partners who had ranked them highly, thus demonstrating a preference for sharing with them. However, participants also sent more money to partners who had frequently matched with them, even though partners controlled only rankings. Moreover, participants who relied more on outcome feedback during learning also relied more on outcome feedback during prosocial decisions. Patterns of instrumental learning from social rejection thus gave rise not only to patterns of partner choice but also to patterns of prosocial behavior within interactions.

General Discussion

Humans face two learning challenges when choosing which social bonds to build. First, people must learn whether others have desirable qualities—for instance, whether others are generous, competent, or cooperative—within a "marketplace" of potential partners (Martin et al., 2019). Second, people must learn whether others see them as desirable partners; unlike goods in other marketplaces, other people evaluate us in return. This metaperception—perceiving whether others value us—generates a unique source of value in partner choice (Byrne & Rhamey, 1965; Heider, 1946; Montoya & Horton, 2014; Shanteau & Nagy, 1979). Whereas past research has examined how people learn to interact with partners who have valuable qualities (Barclay & Willer, 2007; Hackel et al., 2015; Martin et al., 2019), here, we demonstrate that people learn to affiliate with others who value them in part by making choices and experiencing acceptance feedback along two dimensions—outcome and intention. People learn to interact with individuals who show a desire to interact with them and with individuals who do concretely interact with them.

These findings inform how feedback-based learning shapes social interaction above and beyond more passive forms of social experience. Past research has characterized the painful impact of social rejection across the brain (Eisenberger et al., 2003; Kross et al., 2011; Somerville et al., 2006; Woo et al., 2014), affect (DeWall & Bushman, 2011; Leary & Acosta, 2018; Williams et al., 2000), and behavior (Maner et al., 2007; Twenge et al., 2001). In addition, past research has demonstrated that people passively learn cues that predict acceptance or rejection via classical conditioning (Jones et al., 2011; Olsson et al., 2013) and that people gain or lose self-esteem when they receive more or less social approval than expected (Will et al., 2017). Yet, people must also use social feedback to adjust their partner choices, transforming feedback into value representations that guide choice. We provide evidence that people learn to choose partners based on both acceptance outcomes and acceptance intentions following choices, updating representations of a partner's preferences and representations of the likelihood of acceptance.

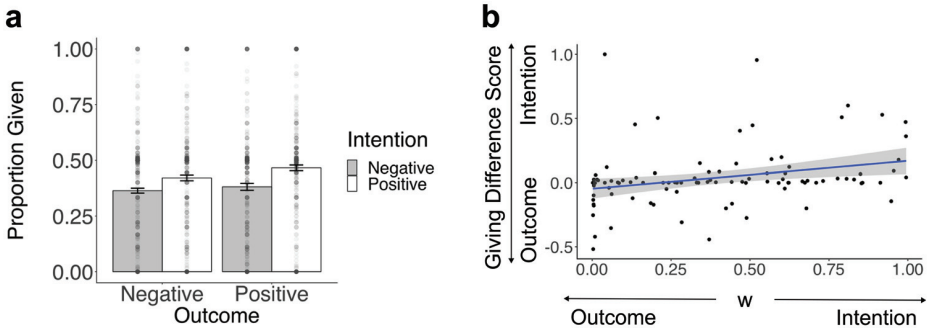
Implications for Instrumental Learning

The present findings offer new insights into the role of active instrumental learning in social affiliation, complementing more passive forms of learning (Amodio, 2019; FeldmanHall et al., 2021; FeldmanHall & Dunsmoor, 2019; Jones et al., 2011; Murty et al., 2016). Models of instrumental learning predict that concretely rewarding outcomes incrementally influence learning and that such learning can give rise to patterns of persistent choice (Balleine & Dickinson, 1998; Gillan et al., 2015; Hackel et al., 2019; Wood & Rünger, 2016). Here, we found that socioemotional outcomes of acceptance serve this role, beyond monetary reward or pleasant social stimuli frequently found to reinforce choice (Hackel et al., 2015; Lin et al., 2012; Lindström et al., 2014): Participants felt better when accepted than rejected and gravitated toward partners who provided that acceptance. Further, participants persisted in choosing those partners even after contingencies changed to render prior outcomes irrelevant. Specifically, during the learning phase of our experiments, some partners were systematically allowed many matches and therefore matched with participants despite ranking the participant poorly; in the test phase, these partners no longer had additional matches available,

[1] Visual inspection of the scatterplot revealed outliers (Figure 4b). We therefore supplemented the analysis with robust regression using the R package MASS (Venables & Ripley, 2002), which supported the initial results, $b = .16$, $SE = .04$, $t(93) = 3.80$, $p < .001$, 95% CI [.08, .24].

Figure 4

Prosocial Giving Based on Learning From Acceptance Intentions and Outcomes in Study 3



Note. (a) The y axis shows the proportion of points participants sent to a partner based on the average rank (intention) and average proportion of matching (outcome) the partner provided during the learning phase. Dots indicate individual data points, with darker shade indicating a higher density of points. Error bars indicate standard error of the mean, with within-participant adjustment (Morey, 2008). Participants gave proportionally more to partners who gave them high average ranks and who were associated with high match frequency. (b) Relationship between reliance on intentions versus outcomes during learning and during the prosocial choice phase in Study 3. The x axis shows a participant's w parameter, indicating whether they relied more on intention (higher values) or outcome (lower values) when choosing partners during learning. The y axis shows an analogous difference score indicating the extent to which a participant repaid money to partners based on the partner's intentions (higher values) versus outcomes (lower values). Shaded region indicates 95% confidence interval. Participants who relied more on outcomes (versus intentions) when choosing partners during the learning phase also relied more on outcomes (versus intentions) when repaying money in the trust game. See the online article for the color version of this figure.

rendering the earlier feedback irrelevant. Nonetheless, participants still preferred interacting with these individuals relative to those who had provided fewer matching outcomes. Participants demonstrated this preference even when social rejection carried no economic consequences. This persistence raises the possibility that people might form relatively habitual tendencies of social interaction through experiencing acceptance feedback (Wood, 2017)—a possibility worthy of future investigation. Altogether, these findings expand models of instrumental learning and their consequences to socioemotional outcomes.

Strikingly, participants also inferred they were better liked by partners who provided positive outcomes, even when those outcomes were explicitly determined by situational constraints rather than the partner's desire to interact with them. Although participants did primarily believe they were liked by those who ranked them highly, they secondarily believed that they were better liked by those who provided more frequent acceptance outcomes. Participants even acted kindlier to these partners, sharing a larger proportion of available money with them. These findings are consistent with prior reports that reward-based learning in social interaction leads people to like others (Hackel et al., 2019, 2020) and to reciprocate with others (Hackel & Zaki, 2018). This influence of reward may reflect affective associations that color social perception. For instance, in the present studies, participants reported feeling more positive affect when accepted than rejected, regardless of a partner's intentions. Participants may have associated this affect with a partner and misattributed it to the partner's intentions. These findings highlight a link between socioemotional reward and social perception.

At the same time, it is also true that social behavior depends on more than the reward cues emphasized in traditional reinforcement learning models; people use models of others' mental states to understand and predict their actions (Tamir & Thornton, 2018; Vélez & Gweon, 2021). In particular, intentions offer an abstract form of learning, indicating whether a partner is likely to value us across varying settings and situational constraints (Kalkstein et al., 2018, 2020; Leary & Acosta, 2018; Trope & Liberman, 2010). For instance, a friend who cares for us might be expected to attend a birthday party, validate our feelings, or sit with us at lunch—scenarios that vary in their specific features but exemplify the abstract concept of care. Moreover, that friend may be unable to do so when looking after a sick parent, reflecting a situational constraint that should not alter one's view of the relationship. This type of abstract learning may be adaptive in maintaining relationships; if a friend is unavailable due to looking after a sick parent, retaliation would be ill-advised. Indeed, participants strongly learned from intention feedback, choosing partners who ranked them positively and inferring they were liked primarily from this feedback. This finding is consistent with social–cognitive models and supports a role for mental state inference in social reinforcement learning.

Although both outcomes and intentions reinforced choice, participants had higher learning rates for intentions than outcomes, indicating they updated intention representations more quickly to align with recent (as opposed to earlier) feedback. These higher learning rates might reflect greater sensitivity to intentions, which reveal one's relational value in the eyes of a partner; greater reliance on working memory, which tends to support faster learning relative to incremental reinforcement processes (Collins et al., 2017; Frank et al., 2007) and which can support goal-directed action planning (Otto et al., 2013); or other differences in learning from probabilistic versus continuous feedback. These possibilities

offer intriguing directions for future work. Nonetheless, these findings highlight the dissociation between these two types of learning, each of which contributed to choice.

Insights Into Social Behavior

By demonstrating that both types of feedback contribute to learning, the present research suggests an expansive role of outcome and intention in social behavior. In Western cultures, people tend to judge others' morality based on both the intentions behind their actions and the outcomes they cause (Martin & Cushman, 2015; Young et al., 2007, 2010). The present research demonstrates that outcome and intention shape not only judgments of morality but also responses to social rejection across affect, partner choice, and social perception. In particular, the present experiments explicitly presented participants with information about a partner's intentions: participants saw exactly how much partners wanted to interact with them relative to others. Nonetheless, participants preferred partners with whom they had previously matched as opposed to partners with whom they did not, even when these partners equally desired to interact with them. These findings held true both when rejection barred material opportunities (Studies 1 and 3) and when rejection signaled only social disregard (Study 2), suggesting that intentions and outcomes influence multiple domains of social behavior. Moreover, we did not observe consistent statistical interactions between these variables,[2] which suggests that that they may exert relatively independent influences on partner choice.

By parameterizing the development of social affiliation and parsing underlying processes, the present model may also support new insights into maladaptive social functioning, which often involves atypical learning from social feedback (Beltzer et al., 2019; Frey et al., 2021; Lamba et al., 2020; Siegel et al., 2020). For instance, we found that people acted kindly or retaliated based on acceptance and rejection outcomes even when outcomes were distinct from intentions, especially among individuals who had relied more strongly on outcome feedback when choosing partners during learning. This tendency could give rise to unwarranted social conflict if people retaliate against rejection that does not reflect another person's preferences. Given that some forms of psychopathology involve strong emotional responses, altered theory of mind, and interpersonal conflict (Beltzer et al., 2019; Berenson et al., 2011; Dixon-Gordon et al., 2015, 2018; Domsalla et al., 2014; Morrison & Heimberg, 2013; Sadikaj et al., 2010; Stepp et al., 2009), dissociating these components of social learning may offer insight into interpersonal difficulties. For instance, interpersonal difficulties might arise from enhanced emotional responses to outcomes, a failure to accurately infer intentions, or both—and these possibilities suggest different targets of intervention.

By explicitly dissociating intentions from outcomes, the present work allowed a strong experimental test of responses to outcomes as well as formal modeling of learning. This approach mirrors common scenarios in which outcomes and intentions diverge, such as knowing one was picked first or last for a team or discovering that one was a first- or second-choice candidate for a job. In other situations, however, intentions must be inferred from ambiguous cues, such as receiving no response to an email. In those cases, intentions and outcomes may covary, and outcomes may have an even stronger influence on behavior. Future work can complement the current experimental approach, which dissociates underlying cognitive processes, by examining how people respond in daily life contexts in which intentions and outcomes cannot be clearly disentangled. In addition, the present experiments involved learning about rejection from strangers—a type of rejection that can sting sharply (Snapp & Leary, 2001). This design mirrored initial stages of relationship development, allowed formal modeling of learning, and avoided preexisting knowledge about others. Future work can test whether people learn differently in preexisting relationships, perhaps by generating strong priors about the intentions of friends or enemies (Kim et al., 2020; Snapp & Leary, 2001). Finally, participants in the present study belonged to WEIRD populations (Henrich et al., 2010), and the extent to which people rely on intentions versus outcomes in moral judgments may vary across cultures (McNamara et al., 2019). Patterns of social learning may similarly vary across cultures.

More broadly, however, the present findings support a hybrid model of social instrumental learning, in which learning across different kinds of social interactions depends on both concretely rewarding outcomes and more abstract social value. For instance, when people learn about social partners who share money, they gravitate toward partners who provide material reward (sharing large amounts of money) and who display generous character (sharing large proportions of available money; Hackel et al., 2015, 2020). These types of learning are associated with neural responses linked to both reward processing and social impression formation, respectively, suggesting a confluence of reinforcement and social cognition (Hackel et al., 2015). The present work analogously demonstrates that people learn which partners value them in return by experiencing socioemotional outcomes and by learning that others want to interact with them. People thus solve two key social learning challenges—learning who to value and learning who values them—through the integration of reward processing and social cognition.

Conclusions

Altogether, the present research highlights the promise of instrumental learning approaches for characterizing socioemotional processes. Formal models of learning have offered insights into feedback-based social behavior, informing the dynamics of learning in active interactions and suggesting links between social behavior and neural computation (Amodio, 2019; Bellucci & Park, 2020; Crockett, 2016; FeldmanHall & Dunsmoor, 2019; Hackel & Amodio, 2018; Hackel et al., 2019; Hertz, 2021; Lockwood & Klein-Flügge, 2020; Kozakevich Arbel et al., 2021; Olsson et al., 2020; Suzuki & O'Doherty, 2020). The current evidence addresses how people learn to approach or avoid others through socioemotional feedback that requires social cognition, thus illuminating computations that underlie rich forms of learning and choice people experience in their social lives.

Context of the Work

How do people learn about others by making choices and experiencing feedback during interaction? Models of reinforcement

[2] While significant interactions were detected in Study 1 affect ratings and Study 3 social perceptions, these interaction effects did not replicate across studies.

learning focus on how people learn from rewarding outcomes (Daw et al., 2011), whereas models of social cognition traditionally focus on how people form inferences about other people's traits and mental states (Uleman & Kressel, 2013). Recent work indicates that neither approach can explain human social behavior alone, consistent with the view that multiple processes of learning and memory contribute to social behavior (Amodio, 2019). In particular, past work explored how people learn which partners to value through monetary reward feedback and feedback indicating others' character traits (e.g., generosity; Hackel et al., 2015). We developed the present work to understand how these processes relate to people's socioemotional learning of which partners value them in return. We proposed, and found across three studies, that acceptance outcomes and cues to acceptance intentions guide learning and downstream behaviors. These findings further illuminate how social learning through feedback can support social behavior.

References

Amodio, D. M. (2019). Social Cognition 2.0: An interactive memory systems account. *Trends in Cognitive Sciences*, *23*(1), 21–33. https://doi.org/10.1016/j.tics.2018.10.002

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology*, *37*(4-5), 407–419. https://doi.org/10.1016/S0028-3908(98)00033-1

Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings. Biological Sciences*, *274*(1610), 749–753. https://doi.org/10.1098/rspb.2006.0209

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *ArXiv*. https://arxiv.org/abs/1406.5823

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529.

Bellucci, G., & Park, S. Q. (2020). Honesty biases trustworthiness impressions. *Journal of Experimental Psychology: General*, *149*(8), 1567.

Beltzer, M. L., Adams, S., Beling, P. A., & Teachman, B. A. (2019). Social anxiety and dynamic social reinforcement learning in a volatile environment. *Clinical Psychological Science*, *7*(6), 1372–1388. https://doi.org/10.1177/2167702619858425

Berenson, K. R., Downey, G., Rafaeli, E., Coifman, K. G., & Paquin, N. L. (2011). The rejection-rage contingency in borderline personality disorder. *Journal of Abnormal Psychology*, *120*(3), 681–690. https://doi.org/10.1037/a0023335

Berenson, K. R., Gyurak, A., Ayduk, O., Downey, G., Garner, M. J., Mogg, K., Bradley, B. P., & Pine, D. S. (2009). Rejection sensitivity and disruption of attention by social threat cues. *Journal of Research in Personality*, *43*(6), 1064–1072. https://doi.org/10.1016/j.jrp.2009.07.007

Bhanji, J. P., & Delgado, M. R. (2014). The social brain and reward: Social information processing in the human striatum. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(1), 61–73. https://doi.org/10.1002/wcs.1266

Bourgeois, K. S., & Leary, M. R. (2001). Coping with rejection: Derogating those who choose us last. *Motivation and Emotion*, *25*(2), 101–111. https://doi.org/10.1023/A:1010661825137

Buckley, K. E., Winkel, R. E., & Leary, M. R. (2004). Reactions to acceptance and rejection: Effects of level and sequence of relational evaluation. *Journal of Experimental Social Psychology*, *40*(1), 14–28. https://doi.org/10.1016/S0022-1031(03)00064-7

Byrne, D., & Rhamey, R. (1965). Magnitude of positive and negative reinforcements as a determinant of attraction. *Journal of Personality and Social Psychology*, *2*(6), 884–889. https://doi.org/10.1037/h0022656

Cacioppo, J. T., & Cacioppo, S. (2014). Social relationships and health: The toxic effects of perceived social isolation. *Social and Personality Psychology Compass*, *8*(2), 58–72. https://doi.org/10.1111/spc3.12087

Cho, H. (2021). *Instrumental learning of social affiliation*. Retrieved from: https://osf.io/6e73p/?view_only=d3b048ddbf9743f19f72740abe37cfd5

Collins, A. G. E., Ciullo, B., Frank, M. J., & Badre, D. (2017). Working memory load strengthens reward prediction errors. *The Journal of Neuroscience*, *37*(16), 4332–4342. https://doi.org/10.1523/JNEUROSCI.2700-16.2017

Crockett, M. J. (2016). How formal models can illuminate mechanisms of moral judgment and decision making. *Current Directions in Psychological Science*, *25*(2), 85–90. https://doi.org/10.1177/0963721415624012

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–126. https://doi.org/10.1037/0022-3514.44.1.113

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In M. R. Delgado, E. A. Phelps, & T. W. Robbins (Eds.), *Decision making, affect, and learning: Attention and performance XXIII* (pp. 3–38). Oxford University Press.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Decker, J. H., Lourenco, F. S., Doll, B. B., & Hartley, C. A. (2015). Experiential reward learning outweighs instruction prior to adulthood. *Cognitive, Affective & Behavioral Neuroscience*, *15*(2), 310–320. https://doi.org/10.3758/s13415-014-0332-5

DeWall, C. N., & Bushman, B. J. (2011). Social acceptance and rejection: The sweet and the bitter. *Current Directions in Psychological Science*, *20*(4), 256–260. https://doi.org/10.1177/0963721411417545

DeWall, C. N., & Richman, S. B. (2011). Social exclusion and the desire to reconnect. *Social and Personality Psychology Compass*, *5*(11), 919–932. https://doi.org/10.1111/j.1751-9004.2011.00383.x

Dewall, C. N., Maner, J. K., & Rouby, D. A. (2009). Social exclusion and early-stage interpersonal perception: Selective attention to signs of acceptance. *Journal of Personality and Social Psychology*, *96*(4), 729–741. https://doi.org/10.1037/a0014634

Dixon-Gordon, K. L., Tull, M. T., Hackel, L. M., & Gratz, K. L. (2018). The influence of emotional state on learning from reward and punishment in borderline personality disorder. *Journal of Personality Disorders*, *32*(4), 433–446. https://doi.org/10.1521/pedi_2017_31_299

Dixon-Gordon, K. L., Weiss, N. H., Tull, M. T., DiLillo, D., Messman-Moore, T., & Gratz, K. L. (2015). Characterizing emotional dysfunction in borderline personality, major depression, and their co-occurrence. *Comprehensive Psychiatry*, *62*, 187–203. https://doi.org/10.1016/j.comppsych.2015.07.014

Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D., & Daw, N. D. (2015). Model-based choices involve prospective neural activity. *Nature Neuroscience*, *18*(5), 767–772. https://doi.org/10.1038/nn.3981

Domsalla, M., Koppe, G., Niedtfeld, I., Vollstädt-Klein, S., Schmahl, C., Bohus, M., & Lis, S. (2014). Cerebral processing of social rejection in patients with borderline personality disorder. *Social Cognitive and Affective Neuroscience*, *9*(11), 1789–1797. https://doi.org/10.1093/scan/nst176

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.

Eastwick, P. W., Finkel, E. J., Mochon, D., & Ariely, D. (2007). Selective versus unselective romantic desire: Not all reciprocity is created equal. *Psychological Science*, *18*(4), 317–319.

Edwards, L. J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger, O. (2008). An R2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine*, *27*(29), 6137–6157. https://doi.org/10.1002/sim.3429

Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An FMRI study of social exclusion. *Science*, *302*(5643), 290–292. https://doi.org/10.1126/science.1089134

FeldmanHall, O., & Dunsmoor, J. E. (2019). 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., Montez, D. F., Phelps, E. A., Davachi, L., & Murty, V. P. (2021). Hippocampus guides adaptive learning during dynamic social interactions. *The Journal of Neuroscience*, *41*(6), 1340–1348. https://doi.org/10.1523/JNEUROSCI.0873-20.2020

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences*, *104*(41), 16311–16316.

Frey, A.-L., Frank, M. J., & McCabe, C. (2021). Social reinforcement learning as a predictor of real-life experiences in individuals with high and low depressive symptomatology. *Psychological Medicine*, *51*(3), 408–415. https://doi.org/10.1017/S0033291719003222

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *37*(7), 1297–1310. https://doi.org/10.1016/j.neubiorev.2013.03.023

Gershman, S. J. (2016). Empirical priors for reinforcement learning models. *Journal of Mathematical Psychology*, *71*, 1–6. https://doi.org/10.1016/j.jmp.2016.01.006

Gillan, C. M., Otto, A. R., Phelps, E. A., & Daw, N. D. (2015). Model-based learning protects against forming habits. *Cognitive, Affective & Behavioral Neuroscience*, *15*(3), 523–536. https://doi.org/10.3758/s13415-015-0347-6

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*(2), 348–362. https://doi.org/10.1037/0022-3514.85.2.348

Hackel, L. M., & Amodio, D. M. (2018). Computational neuroscience approaches to social cognition. *Current Opinion in Psychology*, *24*, 92–97. https://doi.org/10.1016/j.copsyc.2018.09.001

Hackel, L. M., Berg, J. J., Lindström, B. R., & Amodio, D. M. (2019). Model-based and model-free social cognition: Investigating the role of habit in social attitude formation and choice. *Frontiers in Psychology*, *10*, 2592. https://doi.org/10.3389/fpsyg.2019.02592

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

Hackel, L. M., Mende-Siedlecki, P., & Amodio, D. M. (2020). Reinforcement learning in social interaction: The distinguishing role of trait inference. *Journal of Experimental Social Psychology*, *88*, 103948. https://doi.org/10.1016/j.jesp.2019.103948

Hackel, L. M., & Zaki, J. (2018). Propagation of economic inequality through reciprocity and reputation. *Psychological Science*, *29*(4), 604–613. https://doi.org/10.1177/0956797617741720

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(18), 6741–6746. https://doi.org/10.1073/pnas.0711099105

Heider, F. (1946). Attitudes and cognitive organization. *The Journal of Psychology*, *21*(1), 107–112. https://doi.org/10.1080/00223980.1946.9917275

Heider, F. (1958). *The psychology of interpersonal relations*. Wiley. https://doi.org/10.1037/10628-000

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29. https://doi.org/10.1038/466029a

Hertz, U. (2021). Learning how to behave: Cognitive learning processes account for asymmetries in adaptation to social norms. *Proceedings of the Royal Society B*, *288*(1952), 20210293. https://doi.org/10.1098/rspb.2021.0293

Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: A meta-analytic review. *Perspectives on Psychological Science*, *10*(2), 227–237. https://doi.org/10.1177/1745691614568352

Hughes, B. L., Leong, J. K., Shiv, B., & Zaki, J. (2018). Wanting to like: Motivation influences behavioral and neural responses to social feedback. *BioRxiv*. https://www.biorxiv.org/content/10.1101/300657v1.full

Jaeger, B. (2017). *r2glmm: Computes R squared for mixed (multilevel) models. R package version 0.1.2*. https://CRAN.R-project.org/package=r2glmm

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*(8), 589–604. https://doi.org/10.1016/j.tics.2016.05.011

Joiner, J., Piva, M., Turrin, C., & Chang, S. W. C. (2017). Social learning through prediction error in the brain. *NPJ Science of Learning*, *2*(1), 8. https://doi.org/10.1038/s41539-017-0009-2

Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Libby, V., Glover, G., & Casey, B. J. (2011). Behavioral and neural properties of social reinforcement learning. *The Journal of Neuroscience*, *31*(37), 13039–13045. https://doi.org/10.1523/JNEUROSCI.2972-11.2011

Kalkstein, D. A., Hackel, L. M., & Trope, Y. (2020). Person-centered cognition: The presence of people in a visual scene promotes relational reasoning. *Journal of Experimental Social Psychology*, *90*, 104009. https://doi.org/10.1016/j.jesp.2020.104009

Kalkstein, D. A., Hubbard, A., & Trope, Y. (2018). Expansive and contractive learning experiences: Mental construal and living well. In J. Forgas & R. Baumeister (Eds.), *The social psychology of living well* (pp. 223–236). Routledge. https://doi.org/10.4324/9781351189712-13

Kawachi, I., & Berkman, L. F. (2001). Social ties and mental health. *Journal of Urban Health*, *78*(3), 458–467. https://doi.org/10.1093/jurban/78.3.458

Kim, M., Park, B., & Young, L. (2020). The psychology of motivated versus rational impression updating. *Trends in Cognitive Sciences*, *24*(2), 101–111. https://doi.org/10.1016/j.tics.2019.12.001

Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychological Science*, *28*(9), 1321–1333. https://doi.org/10.1177/0956797617708288

Kozakevich Arbel, E., Shamay-Tsoory, S. G., & Hertz, U. (2021). Adaptive empathy: Empathic response selection as a dynamic, feedback-based learning process. *Frontiers in Psychiatry*, *12*, 706474. https://doi.org/10.3389/fpsyt.2021.706474

Kross, E., Berman, M. G., Mischel, W., Smith, E. E., & Wager, T. D. (2011). Social rejection shares somatosensory representations with physical pain. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(15), 6270–6275. https://doi.org/10.1073/pnas.1102693108

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lamba, A., Frank, M. J., & FeldmanHall, O. (2020). Anxiety impedes adaptive social learning under uncertainty. *Psychological Science*, *31*(5), 592–603. https://doi.org/10.1177/0956797620910993

Leary, M. R. (1999). Making sense of self-esteem. *Current Directions in Psychological Science*, *8*(1), 32–35. https://doi.org/10.1111/1467-8721.00008

Leary, M. R. (2005). Sociometer theory and the pursuit of relational value: Getting to the root of self-esteem. *European Review of Social Psychology*, *16*(1), 75–111. https://doi.org/10.1080/10463280540000007

Leary, M. R., & Acosta, J. (2018). Acceptance, rejection, and the quest for relational value. In A. L. Vangelisti & D. Perlman (Eds.), *The*

Cambridge handbook of personal relationships (2nd ed., pp. 378–390). Cambridge University Press. https://doi.org/10.1017/9781316417867.030

Leary, M. R., Springer, C., Negel, L., Ansell, E., & Evans, K. (1998). The causes, phenomenology, and consequences of hurt feelings. *Journal of Personality and Social Psychology*, *74*(5), 1225–1237. https://doi.org/10.1037/0022-3514.74.5.1225

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13–22. https://doi.org/10.1093/biomet/73.1.13

Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience*, *7*(3), 274–281. https://doi.org/10.1093/scan/nsr006

Lindström, B., Selbing, I., Molapour, T., & Olsson, A. (2014). Racial bias shapes social reinforcement learning. *Psychological Science*, *25*(3), 711–719. https://doi.org/10.1177/0956797613514093

Lockwood, P. L., & Klein-Flügge, M. (2020). Computational modelling of social cognition and behaviour—A reinforcement learning primer. *Social Cognitive and Affective Neuroscience*. Advance online publication. https://doi.org/10.1093/scan/nsaa040

Maner, J. K., DeWall, C. N., Baumeister, R. F., & Schaller, M. (2007). Does social exclusion motivate interpersonal reconnection? Resolving the "porcupine problem". *Journal of Personality and Social Psychology*, *92*(1), 42–55. https://doi.org/10.1037/0022-3514.92.1.42

Martin, J. W., & Cushman, F. (2015). To punish or to leave: Distinct cognitive processes underlie partner control and partner choice behaviors. *PLoS ONE*, *10*(4), e0125193. https://doi.org/10.1371/journal.pone.0125193

Martin, J. W., Young, L., & McAuliffe, K. (2019). *The psychology of partner choice*. https://psyarxiv.com/weqhz/

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, *36*(4), 455–470. https://doi.org/10.1016/S0005-7967(97)10031-6

McNamara, R. A., Willard, A. K., Norenzayan, A., & Henrich, J. (2019). Weighing outcome vs. intent across societies: How cultural models of mind shape moral reasoning. *Cognition*, *182*, 95–108. https://doi.org/10.1016/j.cognition.2018.09.008

Mende-Siedlecki, P. (2018). Changing our minds: The neural bases of dynamic impression updating. *Current Opinion in Psychology*, *24*, 72–76. https://doi.org/10.1016/j.copsyc.2018.08.007

Mende-Siedlecki, P., Baron, S. G., & Todorov, A. (2013). Diagnostic value underlies asymmetric updating of impressions in the morality and ability domains. *The Journal of Neuroscience*, *33*(50), 19406–19415. https://doi.org/10.1523/JNEUROSCI.2334-13.2013

Mende-Siedlecki, P., Cai, Y., & Todorov, A. (2013). The neural dynamics of updating person impressions. *Social Cognitive and Affective Neuroscience*, *8*(6), 623–631. https://doi.org/10.1093/scan/nss040

Miller, K. J., Shenhav, A., & Ludvig, E. A. (2019). Habits without values. *Psychological Review*, *126*(2), 292–311. https://doi.org/10.1037/rev0000120

Montoya, R. M., & Horton, R. S. (2014). A two-dimensional model for the study of interpersonal attraction. *Personality and Social Psychology Review*, *18*(1), 59–86. https://doi.org/10.1177/1088868313501887

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Morrison, A. S., & Heimberg, R. G. (2013). Social anxiety and social anxiety disorder. *Annual Review of Clinical Psychology*, *9*, 249–274. https://doi.org/10.1146/annurev-clinpsy-050212-185631

Murray, S. L., & Holmes, J. G. (2009). The architecture of interdependent minds: A motivation-management theory of mutual responsiveness. *Psychological Review*, *116*(4), 908–928.

Murray, S. L., Holmes, J. G., & Collins, N. L. (2006). Optimizing assurance: The risk regulation system in relationships. *Psychological Bulletin*, *132*(5), 641–666.

Murray, S. L., Leder, S., MacGregor, J. C., Holmes, J. G., Pinkus, R. T., & Harris, B. (2009). Becoming irreplaceable: How comparisons to the partner's alternatives differentially affect low and high self-esteem people. *Journal of Experimental Social Psychology*, *45*(6), 1180–1191.

Murty, V. P., FeldmanHall, O., Hunter, L. E., Phelps, E. A., & Davachi, L. (2016). Episodic memories predict adaptive value-based decision-making. *Journal of Experimental Psychology: General*, *145*(5), 548–558. https://doi.org/10.1037/xge0000158

Olsson, A., Carmona, S., Downey, G., Bolger, N., & Ochsner, K. N. (2013). Learning biases underlying individual differences in sensitivity to social rejection. *Emotion*, *13*(4), 616–621. https://doi.org/10.1037/a0033150

Olsson, A., Knapska, E., & Lindström, B. (2020). The neural and computational systems of social learning. *Nature Reviews Neuroscience*, *21*(4), 197–212. https://doi.org/10.1038/s41583-020-0276-4

Otto, A. R., Gershman, S. J., Markman, A. B., & Daw, N. D. (2013). The curse of planning: Dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychological Science*, *24*(5), 751–761. https://doi.org/10.1177/0956797612463080

Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, *124*(1), 54–74. https://doi.org/10.1037/0033-2909.124.1.54

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Palminteri, S., Wyart, V., & Koechlin, E. (2017). The importance of falsification in computational cognitive modeling. *Trends in Cognitive Sciences*, *21*(6), 425–433. https://doi.org/10.1016/j.tics.2017.03.011

Powers, K. E., Somerville, L. H., Kelley, W. M., & Heatherton, T. F. (2013). Rejection sensitivity polarizes striatal-medial prefrontal activity when anticipating social feedback. *Journal of Cognitive Neuroscience*, *25*(11), 1887–1895. https://doi.org/10.1162/jocn_a_00446

Sadikaj, G., Russell, J. J., Moskowitz, D. S., & Paris, J. (2010). Affect dysregulation in individuals with borderline personality disorder: Persistence and interpersonal triggers. *Journal of Personality Assessment*, *92*(6), 490–500. https://doi.org/10.1080/00223891.2010.513287

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*(3), 513–523. https://doi.org/10.1037/0022-3514.45.3.513

Shanteau, J., & Nagy, G. F. (1979). Probability of acceptance in dating choice. *Journal of Personality and Social Psychology*, *37*(4), 522–533. https://doi.org/10.1037/0022-3514.37.4.522

Siegel, J. Z., Curwell-Parry, O., Pearce, S., Saunders, K. E. A., & Crockett, M. J. (2020). A computational phenotype of disrupted moral inference in borderline personality disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*(12), 1134–1141.

Snapp, C. M., & Leary, M. R. (2001). Hurt feelings among new acquaintances: Moderating effects of interpersonal familiarity. *Journal of Social and Personal Relationships*, *18*(3), 315–326. https://doi.org/10.1177/0265407501183001

Somerville, L. H., Heatherton, T. F., & Kelley, W. M. (2006). Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nature Neuroscience*, *9*(8), 1007–1008. https://doi.org/10.1038/nn1728

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *NeuroImage*, *46*(4), 1004–1017. https://doi.org/10.1016/j.neuroimage.2009.03.025

Stepp, S. D., Pilkonis, P. A., Yaggi, K. E., Morse, J. Q., & Feske, U. (2009). Interpersonal and emotional experiences of social interactions in borderline personality disorder. *Journal of Nervous and Mental Disease*, *197*(7), 484–491. https://doi.org/10.1097/NMD.0b013e3181aad2e7

Sun, S., & Yu, R. (2014). The feedback related negativity encodes both social rejection and explicit social expectancy violation. *Frontiers in Human Neuroscience*, *8*, 556. https://doi.org/10.3389/fnhum.2014.00556

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Suzuki, S., & O'Doherty, J. P. (2020). Breaking human social decision making into multiple components and then putting them together again. *Cortex*, *127*, 221–230. https://doi.org/10.1016/j.cortex.2020.02.014

Suzuki, S., Harasawa, N., Ueno, K., Gardner, J. L., Ichinohe, N., Haruno, M., Cheng, K., & Nakahara, H. (2012). Learning to simulate others' decisions. *Neuron*, *74*(6), 1125–1137. https://doi.org/10.1016/j.neuron.2012.04.030

Tamir, D. I., & Hughes, B. L. (2018). Social Rewards: From basic social building blocks to complex social behavior. *Perspectives on Psychological Science*, *13*(6), 700–717. https://doi.org/10.1177/1745691618776263

Tamir, D. I., & Thornton, M. A. (2018). Modeling the predictive social mind. *Trends in Cognitive Sciences*, *22*(3), 201–212. https://doi.org/10.1016/j.tics.2017.12.005

Thorndike, E. (1911). *Animal intelligence*. Hafner.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463. https://doi.org/10.1037/a0018963

Twenge, J. M., Baumeister, R. F., Tice, D. M., & Stucke, T. S. (2001). If you can't join them, beat them: Effects of social exclusion on aggressive behavior. *Journal of Personality and Social Psychology*, *81*(6), 1058–1069. https://doi.org/10.1037/0022-3514.81.6.1058

Uleman, J. S., & Kressel, L. M. (2013). A brief history of theory and research on impression formation. In D. E. Carlston (Ed.), *Oxford library of psychology. The Oxford handbook of social cognition* (pp. 53–73). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199730018.013.0004

Vélez, N., & Gweon, H. (2021). Learning from other minds: An optimistic critique of reinforcement learning models of social learning. *Current Opinion in Behavioral Sciences*, *38*, 110–115. https://doi.org/10.1016/j.cobeha.2021.01.006

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S (fourth edition)*. Springer. https://doi.org/10.1007/978-0-387-21706-2

Will, G. J., Rutledge, R. B., Moutoussis, M., & Dolan, R. J. (2017). Neural and computational processes underlying dynamic changes in self-esteem. *eLife*, *6*, e28098. https://doi.org/10.7554/eLife.28098

Williams, K. D., Cheung, C. K., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the Internet. *Journal of Personality and Social Psychology*, *79*(5), 748–762. https://doi.org/10.1037/0022-3514.79.5.748

Woo, C.-W., Koban, L., Kross, E., Lindquist, M. A., Banich, M. T., Ruzic, L., Andrews-Hanna, J. R., & Wager, T. D. (2014). Separate neural representations for physical pain and social rejection. *Nature Communications*, *5*(1), 5380. https://doi.org/10.1038/ncomms6380

Wood, W. (2017). Habit in personality and social psychology. *Personality and Social Psychology Review*, *21*(4), 389–403. https://doi.org/10.1177/1088868317720362

Wood, W., & Rünger, D. (2016). Psychology of habit. *Annual Review of Psychology*, *67*, 289–314. https://doi.org/10.1146/annurev-psych-122414-033417

Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(15), 6753–6758. https://doi.org/10.1073/pnas.0914826107

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(20), 8235–8240. https://doi.org/10.1073/pnas.0701408104

Zhu, L., Mathewson, K. E., & Hsu, M. (2012). Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(5), 1419–1424. https://doi.org/10.1073/pnas.1116783109

Received August 21, 2021
Revision received December 26, 2021
Accepted December 28, 2021 ■