Journal of Experimental Social Psychology 99 (2022) 104267

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Group value learned through interactions with members: A reinforcement learning account☆

Leor M. Hackel [a,1,*], Drew Kogon [a,1], David M. Amodio [b,c], Wendy Wood [a]

[a] University of Southern California, United States of America
[b] New York University, United States of America
[c] University of Amsterdam, the Netherlands

ARTICLE INFO

Editor: Professor. Rachel Barkan

Keywords:
Prejudice
Instrumental learning
Reward
Attitudes
Habit

ABSTRACT

How do group-based interaction tendencies form through encounters with individual group members? In four experiments, in which participants interacted with group members in a reinforcement learning task presented as a money sharing game, participants formed instrumental reward associations with individual group members through direct interaction and feedback. Results revealed that individual-level reward learning generalized to a group-based representation, as indicated in self-reported group attitudes, trait impressions, and the tendency to choose subsequent interactions with novel members of the group. Experiments 3 and 4 further demonstrated that group-based reward effects on interaction choices persisted even when past group reward value was no longer predicted of future positive outcomes, consistent with a habit-like expression of group bias. These results demonstrate a novel process of prejudice formation based on instrumental reward learning from direct interactions with individual group members. We discuss implications for existing theories of prejudice, the role of habit in intergroup bias, and intervention strategies to reduce prejudice.

When we interact with another person, we form attitudes and interaction patterns based on feedback they provide in the social exchange (Hackel, Doll, & Amodio, 2015; Lott & Lott, 1974). For instance, if another person shares resources with us, then this rewarding experience may lead us to like them (Hackel, Berg, Lindström, & Amodio, 2019), choose to interact with them again (Hackel, Mende-Siedlecki, & Amodio, 2020), and reciprocate with them (Hackel & Zaki, 2018). These interaction patterns are rooted in instrumental learning—a form of learning through reward reinforcement. To date, instrumental learning about others has been explored primarily in the context of one-on-one social interactions.

The individuals we interact with, however, are often associated with social groups. Thus, it is possible that the patterns formed through reinforcement in individual interactions generalize to the value placed on their groups. This process suggests a novel mechanism of group-level prejudice and discrimination that arises from social contact with individuals. The present research examines this mode of learning about groups and explores its implications for subsequent group-based interactions.

1. Learning value through interaction: The role of instrumental learning

Traditional models of impression formation focus on inferences people form about a partner's character traits when learning about their behavior (Heider, 1958; Jones, 1985; Uleman & Kressel, 2013), and traditional research on attitudes has similarly examined how people form attitudes when learning about another person's positive or negative actions (e.g., finding out that someone acted generously or selfishly; Rydell & McConnell, 2006). In contrast, instrumental learning through direct interaction occurs when people perform actions toward another person and experience rewarding feedback. This distinction is theoretically important given extensive evidence of distinct learning and memory systems for these different types of information (Amodio, 2019; Hackel et al., 2019; Wood, 2017). In particular, reward feedback contributes incrementally to the *value* people associate with another person,

☆ This paper has been recommended for acceptance by Professor. Rachel Barkan.

* Corresponding author at: Department of Psychology, University of Southern California, 3620 South McClintock Ave, Los Angeles, CA 90089, United States of America.
E-mail address: lhackel@usc.edu (L.M. Hackel).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jesp.2021.104267
Received 1 January 2021; Received in revised form 2 December 2021; Accepted 3 December 2021
Available online 28 December 2021

or a representation of the anticipated rewards of future interaction. These value representations can guide attitudes and choices to interact with the same person again. Instrumental learning is thus directly linked to approach and avoidance tendencies; rather than prompting people solely to form conceptual associations between a person and a trait or valence, instrumental learning about a person teaches people whether performing different actions will yield positive or negative consequences (Amodio, 2019; Amodio & Ratner, 2011; Wood, 2019). This valuation may be reflected in deliberate consideration of the anticipated benefits of actions as well as in more implicit approach and avoidance reactions that involve habit (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Miller, Shenhav, & Ludvig, 2019; Rangel, Camerer, & Montague, 2008).

In social interactions, instrumental learning can unfold as people experience material rewards (e.g., a gift) or social rewards (e.g., a compliment). In an initial study of how people learn the value of another through reward feedback in direct interaction, Hackel et al.'s (2015) participants played an economic game in which they learned about partners who shared money. Partners varied in both their reward value (indicated by the absolute amount they shared) and in their trait generosity (indicated by the proportion they shared). Participants learned to choose partners who provided large rewards in addition to choosing partners who acted generously. Moreover, this learning was associated with neural activity in the ventral striatum—a region strongly linked to reward-based reinforcement during instrumental learning (Garrison, Erdeniz, & Done, 2013). Finally, participants preferred partners associated with large rewards and subsequently chose to interact with them even when no further economic incentive was available, indicating that reward feedback shaped social value even when independent of others' characteristics (Hackel et al., 2019, 2020). Altogether, this research suggests that people learn to value social partners—discovering whom to approach versus avoid —in part through instrumental learning of reward value during social interactions.

2. Instrumental learning and generalization to groups

To date, social instrumental learning has been explored primarily in interactions with single individuals. Indeed, instrumental learning cannot directly lead people to form a value representation for a group, given that a person typically does not interact with an entire group at once. That is, a person typically usually cannot interact with all members of a group simultaneously, experience feedback, and update a value representation for the group as a whole.

It is possible, however, for a person to generalize instrumental learning from an individual group member to their group as a whole. In this case, they may form group-level value associations that lead them to approach or avoid members of the group in general. People readily perceive others in terms of social categories, ranging from race, ethnicity, and nationality to university affiliation, sports teams, and political parties. Such social categories play a major role in social interaction, shaping behavioral expectancies (Darley & Gross, 1983; Macrae, Bodenhausen, & Milne, 1995) and basic forms of social perception such as visual face encoding and individuation (Hackel, Looser, & Van Bavel, 2014; Kawakami, Amodio, & Hugenberg, 2017; Ratner & Amodio, 2013; Van Bavel, Packer, & Cunningham, 2011). To the extent that people encode individual interaction partners as members of a social group, they may generalize reward feedback from these interactions to the broader group, much as people generalize other forms of learning from an exemplar to a broader category (Dunsmoor & Murphy, 2014) or from an individual to a group (Hertz, 2021). If so, then this pattern of generalization would be evident in one's tendency to approach or avoid previously unencountered members of the same group based on group-level value representations.

This active form of learning differs from more passive learning mechanisms involving observation or instruction studied in past research on attitudes toward social groups. For instance, perceivers form conceptual impressions of another person's traits when witnessing or hearing about someone's behavior, such as when, upon reading that someone gave to charity or aced a test, they are inferred to be kind or competent (Heider, 1958; Winter & Uleman, 1984). These impressions can be generalized to a group, such that people associate a trait with a group as a whole rather than with individuals alone (Crawford, Sherman, & Hamilton, 2002). Group attitudes may also form passively through the repeated viewing of a social group paired with positive or negative stimuli ("evaluative conditioning;" De Houwer, Thomas, & Baeyens, 2001; Olson & Fazio, 2001). Finally, people may passively learn about groups through propositional processes, in which exposure to information about a group shapes their explicit and implicit group attitudes (De Houwer, 2006; Gregg, Seibt, & Banaji, 2006). Although these passive learning processes provide an important source of attitudes, they do not capture the experience of learning about an individual through the process of receiving reward feedback in direct social interaction.

Instrumental learning, in contrast, involves active learning from the outcomes of social choices; if one receives rewards from past interactions with individual group members, suggesting that their group has high value, then one may pursue future interactions with novel members of that group. Although people often form prejudice in the absence of direct interaction, theory and evidence in cognitive neuroscience give reason to think that interactive learning stems from distinct mechanisms and carries distinct consequences relative to more passive forms of learning (Amodio, 2019; Foerde, Knowlton, & Poldrack, 2006; Hackel et al., 2019; Poldrack et al., 2001; Wood, 2017). That is, whereas passive forms of learning may lead people to apply an attitude or belief to a group-based judgment, instrumental learning has more direct implications for intergroup actions. In this way, instrumental learning augments the extensive research on intergroup contact, identifying an additional mechanism through which direct interactions with group members can influence group-level attitudes and decisions to interact with novel group members, beyond the standard accounts of increasing knowledge about an outgroup, reducing anxiety, and increasing empathy (Pettigrew & Tropp, 2008).

Finally, an instrumental learning perspective suggests a theoretical basis for the notion that people can develop habits of intergroup interaction (Devine, 1989). Repeated instrumental learning can give rise to interaction habits whereby people continue to perform previously rewarded actions without deliberation or intention—even when those actions are no longer relevant to current goals (Wood & Rünger, 2016). For instance, in classic tests of habits, animals that are repeatedly reinforced for pressing a lever for a food reward will continue to do so even after they are no longer hungry or after the reward contingencies have changed (Balleine & Dickinson, 1998). Thus, an instrumental learning account of group-based response formation suggests a mechanism through which people might form habits to interact with members of particular social groups (Wood, 2017, 2019)—a possibility that would not be expected to arise from conceptual, passive forms of learning (Amodio, 2019).

In summary, if people generalize instrumental reward associations from individuals to the social groups those individuals belong to, then they would have a tendency to approach or avoid novel members of those groups. This possibility suggests a yet-unexplored pathway through which people generalize the value associated with group members, rooted in reward feedback and action tendencies. Furthermore, this learned value might give rise to habit-like responding in group-based interaction.

3. Overview

The present research tests whether people generalize reward-based learning from individuals to groups, forming group-level value representations expressed in attitudes and in subsequent choices to interact with novel, previously unencountered group members. Specifically, we conducted a series of studies in which participants iteratively learned

about the rewards provided by interaction partners who belonged to different social groups (students at different universities or members of trivia teams). Afterward, participants completed a test phase in which they could choose to interact with both original and novel members of each group. This procedure allowed us to test whether participants generalized their learning to choices of novel interaction partners. Participants additionally rated their attitudes toward each group and impressions of group members, allowing us to test whether participants formed group-based attitudes and trait inferences through instrumental learning. We hypothesized that participants would generalize instrumental learning from individuals to groups, leading them to choose to interact with novel members of groups that previously provided large rewards.

4. Study 1

In Study 1, participants completed an economic game that was adapted from prior reinforcement learning tasks (Frank, Seeberger, & O'reilly, 2004). In a learning phase, they learned the reward value of interacting with students from four different universities, each of which was associated with a different level of reward. Anonymous university groups were used to avoid any prior stereotypic associations participants might have with existing social groups, allowing us to focus solely on the effects of feedback-based learning.

In a subsequent test phase, participants made additional choices of students, this time without receiving feedback about earnings, in order to assess already-formed associations without new learning. Critically, the test phase featured original and novel members of each group. This manipulation allowed us to test whether participants generalized reward associations with group members to choices of newly encountered group members. We further expected that participants would form more favorable impressions and attitudes toward individuals from more rewarding groups. If these impressions and attitudes were applied to the group as a whole, across original and novel members, then this finding would provide additional evidence of generalization to groups.

4.1. Method

4.1.1. Participants

Fifty-one undergraduate students (20 male, 31 female) participated for class credit or compensation ($10). Sample size was set by aiming for a minimum of 40 participants and then continuing data collection until the end of the semester; this sample size was chosen based on prior research that used similar tasks with a multi-trial within-subjects design (Frank et al., 2004; Hackel et al., 2015). Participants were excluded from analysis if they had extreme response times (+/− 2 SDs from mean), missed more than 10% of responses, and/or pressed the same key more than 90% of the time (Gillan, Otto, Phelps, & Daw, 2015; Hackel et al., 2020; Hackel & Zaki, 2018). These a priori rules identified six participants to exclude, leaving 45 participants in the analyses. Information regarding our procedure for determining sample size, all data exclusions, all manipulations, and all measures included in this research are fully reported in this article. De-identified data from each study have been made available at: https://osf.io/7nyaj/?view_only=fa4aaf673ab14b7f9f0539316fb82fbe.

4.1.2. Procedure

4.1.2.1. Learning phase. Participants first completed 180 learning trials as part of a sharing game; this task was modeled after prior studies of instrumental learning (Frank et al., 2004; Frank, Moustafa, Haughey, Curran, & Hutchison, 2007). In each trial, participants saw a fixation cross (1 s) before viewing images of two individuals represented by avatars, presented side-by-side as a pair (2 s). Avatars represented students from different universities who had ostensibly participated at an earlier time and made a sequence of decisions to share or keep monetary rewards with future participants (Fig. 1a). Upon viewing the avatars, participants selected which of these two students they wished to interact with (Fig. 1b). After each selection, participants received feedback (1 s) indicating whether the chosen student shared a point with them; points were exchanged for money at the end of the study. No feedback was given about the unchosen avatar.

Student avatars were supposedly from one of four universities. University was identified by a red letter and the color of the avatar's shirt. Three avatars from each university were viewed during the learning phase, and an additional three avatars from each university were viewed in the testing phase. In addition, participants were randomly assigned to view all-female or all-male avatars in order to keep gender consistent within subject but allow greater generalizability across subjects. Analyses revealed no effects of participant gender or avatar gender. Thus, these variables are not discussed further.

During the learning phase, points earned on each trial varied with university affiliation. In this phase, participants always viewed pairs of students affiliated with two different universities (AB and CD), following past work using non-social stimuli (Doll, Bath, Daw, & Frank, 2016). In AB trials, choosing an avatar from group A led to a reward on 70% of trials and choosing an avatar from group B led to a reward on 30% of trials. In CD trials, choosing an avatar from group C led to a reward on 60% of trials and choosing an avatar from group D led to a reward on 40% of trials. Different members of a group thus shared at the same rate. University colors were randomly assigned to these roles across participants. Groups of players from each university consisted of six total avatar stimuli, three presented in both learning and test phases (*original avatars*) and three presented only in the test phase (*novel avatars*). Participants thus learned, through instrumental choice and feedback, about the reward value obtained by interacting with members of each group.

4.1.2.2. Test phase. In the subsequent test phase (180 trials), participants made additional choices without receiving feedback, allowing participants to express reward associations in the absence of further learning. Critically, we presented participants with both original and novel faces from each group. Participants again saw AB pairs and CD pairs, but in half these trials, the avatars from each group had been viewed during learning, whereas in the other half of trials, new avatars from each group were presented. (Participants always saw two original avatars or two novel avatars paired on each trial.) In this manner, we examined whether people chose novel group members based on past reward outcomes with other members of the same group. That is, we tested whether participants would choose a novel member of group "A" over a novel member of group "B."

Finally, after the choice task, participants completed three sets of ratings. First, to test whether reward learning also gave rise to explicit impressions, participants rated the generosity of each individual avatar, including both original and novel avatars. Ratings were made on a Likert scale ranging from 1 *(not at all)* to 7 *(very much)*. Second, to examine attitudes toward groups as a whole, participants rated their attitudes toward each university overall using a feeling thermometer scale ranging from 0 (*very cold*) to 100 (*very warm*). Finally, to examine whether results depend on explicit memory for faces, participants also reported whether they recalled seeing each avatar during the learning phase. Analyses revealed that results held across high and low levels of explicit memory, and thus, this variable is not discussed further (see Supplemental Materials for details).

4.2. Results

4.2.1. Test phase

To test our central hypothesis that participants' reward associations with individual group members, learned through direct interaction and feedback, generalized to novel group members, we examined

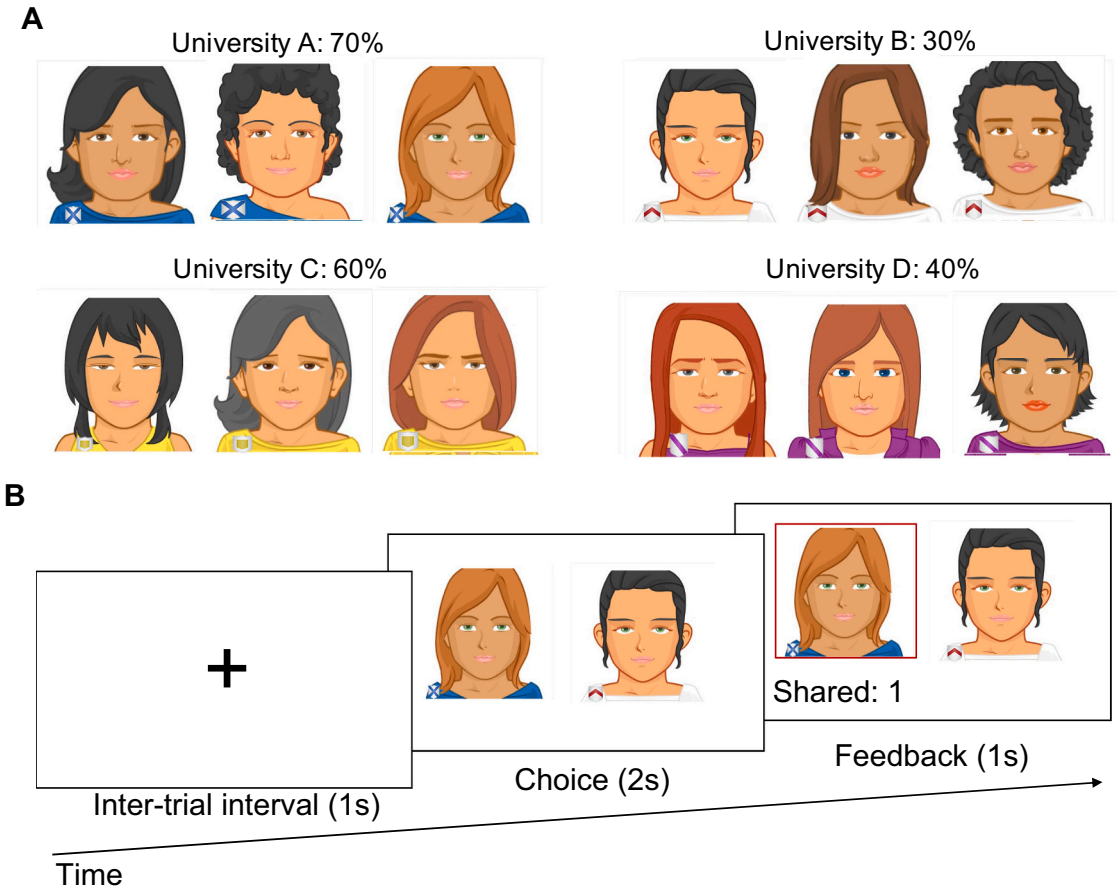

Fig. 1. Schematic of learning task and sample stimuli. (A) Participants learned about ostensible students from four different universities; university affiliation was indicated by shirt color and letters (Study 1) or logos (Studies 2 and 3) on the avatar's shirt. (B) In the learning phase, participants made choices to interact with one of two avatars from different universities on each round. After each choice, participants received feedback indicating whether that player shared one cent out of two.

participant choices in the test phase. As in past work using similar tasks (Hackel et al., 2015), choice of avatars was analyzed using a mixed effects logistic regression. The outcome variable indicated whether participants chose the target (of the two onscreen) from the group that had been more rewarding during learning (1 = *yes,* 0 = *no*). Effect-coded predictors included university pair (A/B = 1, C/D = −1), which indicated the discriminability of reward levels associated with different pairs, and familiarity of face avatars (original = −1, novel = 1). This model therefore revealed whether participants chose members of previously rewarding groups overall (revealed by the intercept), and whether this effect emerged specifically for familiar and for novel members (revealed in simple effects analyses). Data were fit to the model using the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2015; R Core Team, 2016). In all analyses, random variances were included for the intercept and all slopes, nested within subjects. Random effects for avatars were not included, since avatars were randomly assigned to different task roles across participants, avoiding any systematic effects of avatar images on choice.

Overall, participants were likely to choose a member of the previously rewarding group in each pairing, as indicated by an intercept significantly greater than zero, $b = 0.27$, $SE = 0.04$, $z = 7.36$, $p < .001$ (Fig. 2). In addition, a main effect of pair type indicated that participants were more likely to do so for A/B pairs as opposed to C/D pairs, $b = 0.16$, $SE = 0.03$, $z = 6.28$, $p < .001$, consistent with the idea that A/B pairs were easier to discriminate than C/D pairs.

Critically, participants applied reward-based learning to both

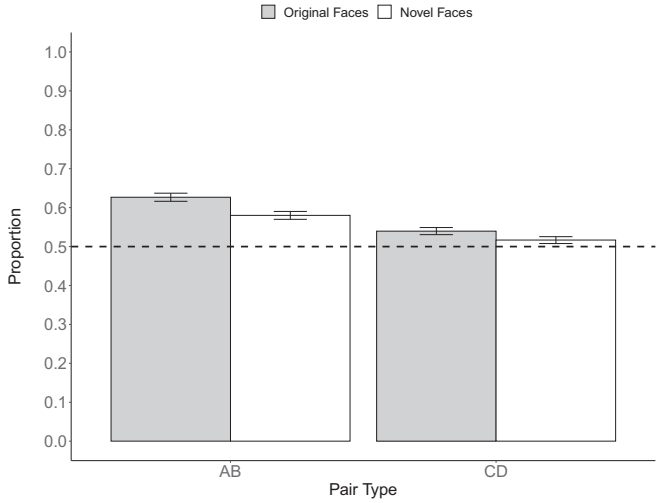

Fig. 2. Test phase choice in Study 1, showing the proportion of trials in which participants chose original and novel members of each group. Participants were more likely to choose members of groups previously associated with higher (as opposed to lower) reward, across original and novel members. The dotted line indicates chance. Error bars show standard error of the mean, with within-participants adjustment (Morey, 2008).

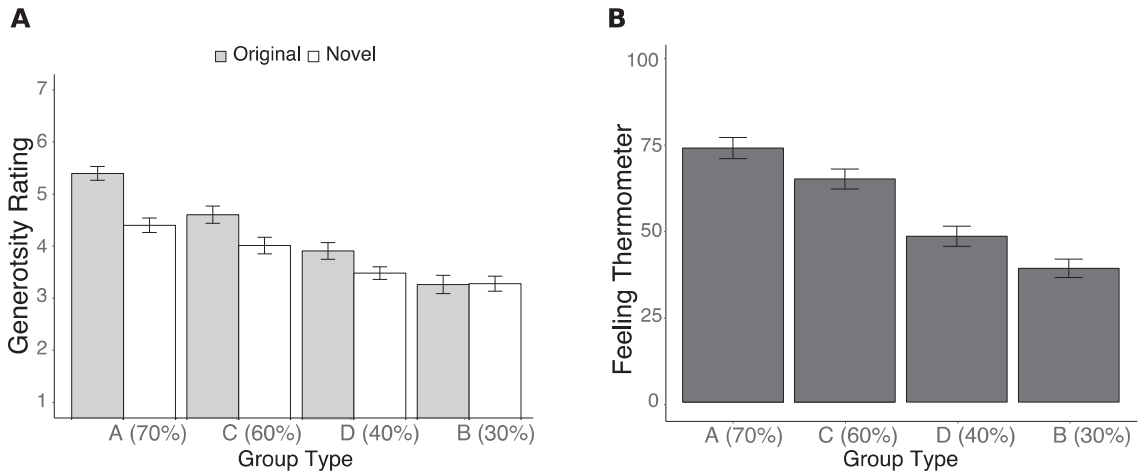

Fig. 3. Explicit ratings of attitudes and impressions in Study 1. (A) Participants rated individual group members as more generous if their groups had been associated with higher reward, across original and novel exemplars. (B) Participants had more positive attitude toward groups as a whole for groups that were associated with higher previous reward.

familiar and novel group members. Simple effects analysis revealed that people chose previously rewarding groups for both familiar faces, $b = 0.34$, $SE = 0.04$, $z = 7.91$, $p < .001$, and novel faces, $b = 0.20$, $SE = 0.04$, $z = 4.58$, $p < .001$, even though reward effects were stronger for familiar avatars (a main effect of stimulus familiarity, $b = -0.07$, $SE = 0.02$, $z = -3.19$, $p = .001$). Thus, participants generalized their reward learning to novel group members. Pair type did not significantly moderate any effects of familiarity, $b = -0.03$, $SE = 0.02$, $z = -1.17$, $p = .24$, indicating that participants relied on prior learning for novel group members to a similar extent across AB pairs and CD pairs. These findings suggest that people generalized reward associations with individuals to a group-level representation, which then guided choices regarding novel group members.

4.2.2. Explicit impressions and attitudes

Next, to determine whether reward learning carried forward into explicit impressions of individual targets and attitudes toward each group, we examined participants' post-task ratings of each avatar's generosity using linear mixed effects regression. Predictors included the reward value of each avatar's group (mean-centered) and familiarity (original vs novel). Analyses were performed using the lme4 and lmerTest packages for R (Bates et al., 2015; Kuznetsova, Brockhoff, & Christensen, 2016; R Core Team, 2016).

This analysis revealed that avatars from more rewarding groups were perceived to be more generous than those from less rewarding groups, $b = 3.83$, $SE = 0.52$, $t(44) = 7.32$, $p < .001$ (Fig. 3a). A main effect of familiarity indicated that original avatars were seen as more generous than novel group members, $b = -0.25$, $SE = 0.06$, $t(44) = -4.47$, $p < .001$, although this main effect was qualified by an interaction with reward value, $b = -1.08$, $SE = 0.29$, $t(224) = -3.67$, $p < .001$. Specifically, group reward value had a stronger impact on impressions of generosity for original (as opposed to novel) avatars, consistent with the pattern of generalization decrement observed in the choice data. That is, although reward value influenced judgments of both original members, $b = 4.91$, SE = 0.60, $t(74.74) = 8.18$, $p < .001$, and novel members, $b = 2.75$, SE = 0.60, t (74.74) = 4.58, $p < .001$, this effect was stronger for original members. Nonetheless, this finding indicates that participants formed impressions of generosity for both original and novel group members based on reward feedback.

Did reward feedback also lead participants to form explicit attitudes toward each group as a whole? To address this question, we examined feeling thermometer ratings toward each group. Ratings were analyzed using linear mixed effects regression, with group reward level as a predictor (mean-centered) and the inclusion of a random intercept and random slope. Given that participants made only one feeling thermometer rating toward each group, this analysis did not include familiarity of faces as a factor. Participants made more favorable ratings of groups that provided more frequent rewards, $b = 87.58$, $SE = 9.78$, t (178) = 8.96, $p < .001$ (Fig. 3b).

Did these explicit attitudes and impressions about groups fully account for participants' choices, or did the effects of reward feedback influence choices independently of these self-reports? To address this question, we refit our regression model predicting test phase choice while accounting for feeling thermometer ratings of each group and generosity ratings for each group. Specifically, we added as a predictor the difference in participants' attitudes toward the two groups onscreen (higher reward group minus lower reward group), along with the interaction of attitudes with face familiarity. We further added the difference in mean generosity ratings toward each group (computed separately for original and novel avatars) as a predictor; given that the means were computed separately for original and novel avatars, these values were not interacted with familiarity. Although explicit impressions of generosity were a significant predictor of choice, $b = 0.11$, $SE =$ 0 0.02, $z = 6.09$, $p < .001$, the intercept remained significantly positive, $b = 0.15$, $SE = 0.03$, $z = 4.50$, $p < .001$, indicating that explicit ratings did not fully account for the impact of reward feedback on choices.[2] These findings suggest that prior reward feedback influenced subsequent choices independent of either explicit attitudes or impressions.

4.3. Discussion

Study 1 revealed that people learn about social groups through generalization of reward-based reinforcement: Through interactions with individual group members, participants formed reward associations with their groups. Furthermore, this learning generalized to choices to interact with novel group members in subsequent encounters. These findings demonstrate that people learn to value social groups based on direct social interactions with individual members.

Reward feedback also led participants to form attitudes and impressions toward groups as a whole: Participants felt warmer toward groups associated with greater reward and rated those groups as being more generous. At the same time, the effect of reward on choice was not fully accounted for by explicit attitudes or impressions, suggesting dissociable influences of reward feedback on choice and explicit

[2] Feeling thermometer ratings did not predict choices over and above the other terms in the model, $b = 0.01$, $SE = 0.03$, $z = 0.45$, $p = .65$.

attitudes. Together, these findings provide initial support for the hypothesis that instrumental reward learning gives rise to group-based partner choice and to group-level attitudes. Participants learned to value interactions with particular groups, shaping their attitudes and choices, via generalization of reward associations.

This finding provides evidence that instrumental learning from interactions with individual group members contributes to the value placed on their group. Moreover, it demonstrates that a group-level reward representation, acquired through interactions with specific individuals, is then generalized to novel members of the group. Although instrumental reward associations with individual group members influenced explicit group attitudes and personality impressions of individual members, it also affected future interaction choices even after adjusting for these explicit attitudes and impressions, suggesting an implicit effect of group-based choice.

5. Study 2

The instrumental associations learned through reward in Study 1 might have influenced subsequent group interaction choices in several ways. Group value could be captured in tendencies to approach more rewarding groups, avoid less rewarding ones, or both. Additionally, the specificity of these value assessments is not clear. Participants might have simply learned to choose one group over another ("always choose Group A over Group B") or they might have formed specific value representations for each group and used these fine-grained distinctions when making choices. Study 2 was designed to distinguish these different types of instrumental associations.

Participants in Study 2 viewed recurring pairings of groups in the learning phase, as in Study 1: AB (70% vs 30%) or CD (60% vs 40%). However, they viewed all possible pairings of the groups in the test phase (i.e., including AC, AD, BC, and BD). These previously unseen pairings, or *transfer* pairings, dissociate the extent to which people learn to approach others through positive feedback as opposed to avoid others through negative feedback. Neural models suggest that separate pathways are involved in processing positive feedback (i.e., reward) and negative feedback (i.e., lack of reward) in this task (Frank et al., 2004, 2007). During learning, people can learn to choose group A over group B either by learning to approach A through positive feedback or by learning to avoid B through negative feedback. Transfer trials dissociate these types of learning: The extent to which people approach A over C and D ("Approach A" trials) reveals positive learning toward A, whereas the extent to which people avoid B in relation to C and D ("Avoid B" trials) reveals negative learning toward B (Frank et al., 2004). By including these transfer trials, we therefore were able to test whether people generalize positive learning, negative learning, or both to novel group members.

Transfer trials also more directly reveal the nature of instrumental learning rooted in basal ganglia function. These pairings require people to transfer value learning by making fine-grained value distinctions (e.g., 70% vs 60%). Performance on stimulus transfer trials correlates with genetic markers of striatal dopamine (Doll et al., 2016) and is particularly susceptible to dopaminergic manipulations (Frank et al., 2004; Jocham, Klein, & Ullsperger, 2011). As a result, performance on transfer pairings may provide an even stronger index of instrumental learning than in Study 1.

5.1. Method

5.1.1. Participants

Eighty undergraduate students (48 female, 32 male) participated. A power analysis using bootstrapped simulations of Study 1 data revealed that 80 participants offered greater than 99% power to detect the simple effect of prior reward feedback when interacting with new stimuli in the test phase. Four participants were excluded from analysis due to failure to meet our inclusion criteria described in Study 1, leaving 76 participants for analysis.

5.1.2. Procedure

The procedure was identical to that of Study 1, with one exception: the test phase of the learning task featured all possible pairings of groups (e.g., A paired with B, C, and D). Each pairing of groups appeared equally often in the test phase. As in Study 1, this task included 180 learning trials and 180 test trials and was followed by a post-task questionnaire. Test phase trials were evenly split between trials featuring original and novel avatars.

5.2. Results

5.2.1. Test phase

Did participants both approach and avoid novel group members on the basis of prior reward feedback? To test this question, choice of avatars was again analyzed using a mixed effects logistic regression designed to predict whether participants chose members of groups associated with greater rewards during the learning phase. We first verified that results from Study 1 replicated when analyzing A/B and C/D trials in the test phase using the same predictors as in Study 1. As anticipated, we observed a significant intercept, $b = 1.26$, $SE = 0.17$, $z = 7.51$, $p < .001$, indicating a greater tendency to choose previously rewarding targets. Simple effects analysis indicated that this was true for original group members, $b = 1.42$, $SE = 0.18$, $z = 7.97$, $p < .001$, and novel group members, $b = 1.11$, $SE = 0.18$, $z = 6.32$, $p < .001$, even though this effect was relatively stronger for original members (a main effect of familiarity, $b = 0–0.15$, $SE = 0.05$, $z = -2.80$, $p = .005$).

Next, consistent with prior work using this task, we analyzed *transfer trials*, or the trials featuring unpracticed pairings, because, as explained above, these trials index instrumental learning (Doll et al., 2016) and dissociate approach and avoidance learning (Frank et al., 2004). Predictors included avatar familiarity (−1 = original, 1 = novel) and approach vs. avoidance learning (−1 = avoid B, 1 = choose A).

Participants chose members of groups associated with high reward value overall, as indicated by a positive intercept, $b = 0.89$, $SE = 0.13$, $z = 7.06$, $p < .001$, but critically, this was true for both original and novel members (Fig. 4). Simple effects analysis revealed that participants relied on prior reward learning both when choosing original avatars, b

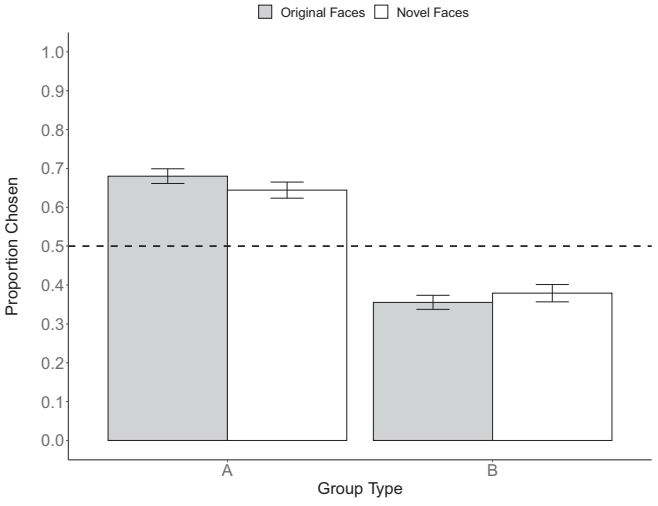

Fig. 4. Test phase choice in Study 2, showing the proportion of trials in which participants chose members of groups previously associated with higher reward value, across choosing Group "A" in "Approach A" trials and Group "B" in "Avoid B" trials. Participants were more likely to choose members of groups previously associated with higher (as opposed to lower) reward, across original and novel members. The dotted line indicates chance. Error bars show standard error of the mean, with within-participants adjustment (Morey, 2008).

$= 0.98$, $SE = 0.13$, $z = 7.43$, $p < .001$, and when choosing novel avatars, $b = 0.81$, $SE = 0.13$, $z = 6.13$, $p < .001$, although the effect of reward was stronger for original avatars (i.e., a main effect of familiarity, $b = -0.09$, $SE = 0.04$, $z = -2.35$, $p = .02$).

In addition, participants chose group members on the basis of approach learning from positive feedback and avoidance learning from negative feedback. We did not observe a significant main effect of approach/avoidance, $b = 0.08$, $SE = 0.09$, $z = 0.88$, $p = .37$, or an interaction with familiarity, $b = -0.03$, $SE = 0.04$, $z = -0.65$, $p = .52$, indicating that participants chose novel avatars based on prior reward learning across "approach A" and "avoid B" trials. That is, across familiar and novel avatars, participants were likely to approach members of Group A over Groups C and D and were likely to avoid members of Group B for members of Groups C and D. This finding indicates that participants acquired reward associations through both positive and negative feedback and expressed these associations toward novel group members.

5.2.2. Explicit impressions and attitudes

Post-task ratings of generosity replicated all findings from Study 1: Participants rated members of rewarding groups as more generous, $b = 3.88$, $SE = 0.45$, $t(75) = 8.69$, $p < .001$ (Fig. 5a). Again, original group members were rated as more generous than novel ones, $b = -0.23$, $SE = 0.04$, $t(454) = -6.04$, $p < .001$, but this effect was moderated by an interaction with group reward value, $b = -0.79$, $SE = 0.24$, $t(454) = -3.30$, $p < .001$. Specifically, participants especially rated members of rewarding groups as more generous when viewing original, as opposed to novel, partners. Nonetheless, as in Study 1, simple effects analysis revealed that reward feedback influenced ratings of both original partners, $b = 4.67$, $SE = 0.51$, $t(123.22) = 9.22$, $p < .001$, and novel partners, $b = 3.08$, $SE = 0.51$, $t(123.22) = 6.08$, $p < .001$. These findings again demonstrate that instrumental learning involving individual group members led people to form explicit impressions of a group's generosity, applied to original and novel members.

In feeling thermometer ratings, participants again made more favorable ratings of groups that shared more often, $b = 80.74$, $SE = 8.95$, $t(75) = 9.02$, $p < .001$ (Fig. 5b). This finding verifies Study 1 in that participants formed attitudes toward groups as a whole based on instrumental reward feedback from its individual members.

Once again, however, explicit attitudes and impressions did not fully account for patterns of reward-based choice. When adding feeling thermometer scores and impressions for each group as a predictor in the analysis of test phase choices on transfer trials, choices were related to both attitudes, $b = 0.11$, $SE = 0.04$, $z = 2.55$, $p = .01$, and impressions, $b = 0.74$, $SE = 0.09$, $z = 0.91$, $p < .001$. Nonetheless, the intercept remained significantly positive, $b = 1.06$, $SE = 0.14$, $z = 7.53$, $p < .001$, indicating that explicit ratings did not fully account for choice behavior. Thus, reward feedback shaped choice in a manner dissociable from its effect on explicit attitudes and impressions.

5.3. Discussion

Study 2 replicated and extended the findings of Study 1: Through interactions with individual group members, participants formed reward associations with their groups and, subsequently, applied this learning to novel group members. Furthermore, Study 2 linked these tendencies more closely to instrumental learning: These choices persisted in unpracticed pairings that required participants to make fine-grained value distinctions, which have been linked to striatal-based instrumental learning in past research. Moreover, these findings held across indicators of positive and negative learning, suggesting that participants similarly learned to approach and avoid members of social groups, with a slightly greater tendency to approach highly rewarding group members than to avoid less rewarding ones.

Finally, instrumental learning again led participants to form explicit impressions of a group's generosity—which were applied to both original and novel group members—as well as explicit group-based attitudes. At the same time, explicit attitudes and impressions again did not fully account for the effect of reward feedback on choices, as in Study 1, further suggesting an implicit influence of reward feedback on behavior. The nature of this direct effect was evaluated in Study 3.

6. Study 3

Study 3 was designed to more directly isolate the role of reward feedback in intergroup interactions and to test its persistence in influencing choice. Studies 1 and 2 suggested that instrumental reward associations may directly shape subsequent interaction choices. In the present study, we aimed to determine the extent to which this direct effect represents an effect of reward feedback as opposed to feedback about a group's traits.

To dissociate the effects of reward feedback and character feedback on choice, we used a learning task that independently manipulates the reward an avatar provides and the generosity an avatar displays (Hackel et al., 2015). This task allowed us to isolate the impact of reward while experimentally controlling trait feedback. On each round, participants interacted with avatars who had a pool of points available and shared a proportion of those points. Some avatars shared a large proportion, on

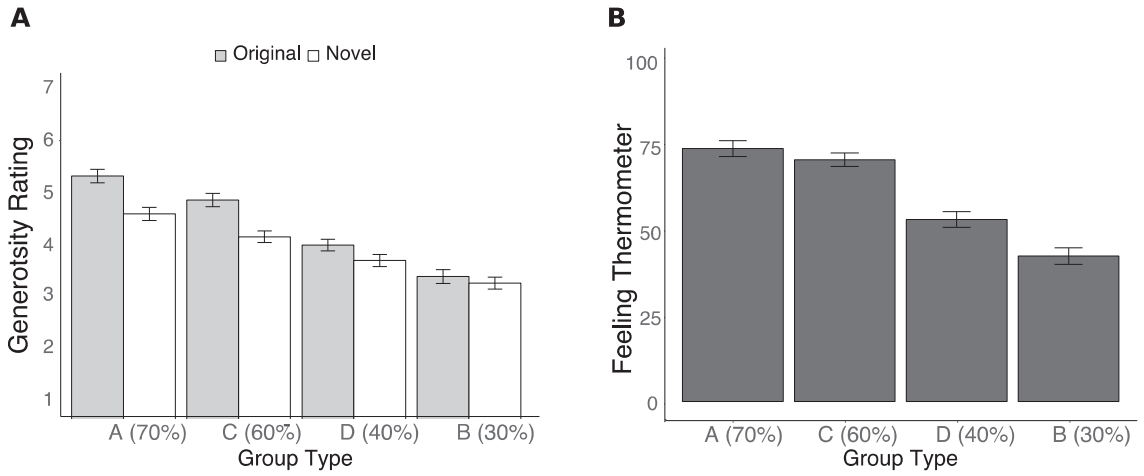

Fig. 5. Explicit ratings of attitudes and impressions in Study 2. (A) Participants rated individual group members as more generous if their groups had been associated with higher reward, across original and novel exemplars. (B) Participants had more positive attitude toward groups as a whole for groups that were associated with higher previous reward.

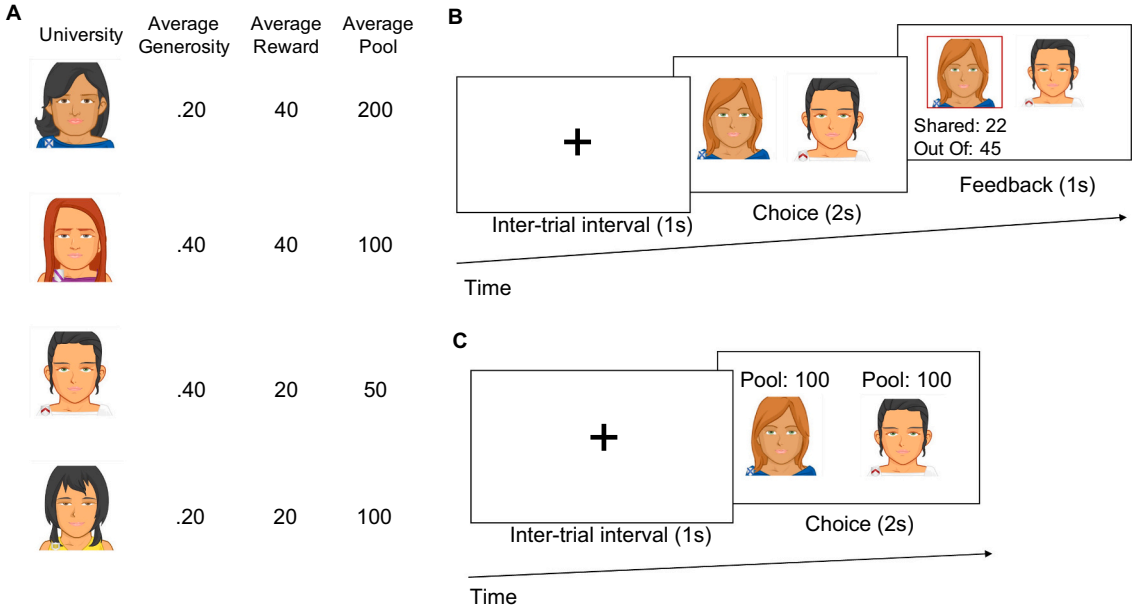

Fig. 6. Schematic of study design in Study 3. (A) Groups varied orthogonally in the average reward they provided (amount shared) and average generosity they displayed (proportion shared). Some groups had larger point pools, on average, rendering reward statistically independent of generosity. (B) In a learning phase, participants made choices to interact with one of two avatars from different universities on each round. After each choice, participants received feedback displaying the amount shared and the point pool that player had available, indicating proportional generosity. (C) In a test phase, participants made further choices without feedback. The point pools available to each player were displayed above avatars, rendering prior reward associations irrelevant.

average, revealing high generosity, and some avatars shared a large number of points, revealing their reward value (Fig. 6a).

This task further allowed us to begin exploring the extent to which reward-based decisions were rooted in goal-directed or relatively habit-like behavior. In test phase trials of this task, the task contingencies were changed: all Deciders now had an equal number of points, and this information was shown to participants when choosing between players. As a result, only a player's prior generosity would now predict their sharing. For instance, if two players have 100 points to share, then the more generous one will offer a larger reward. Although some Deciders had more points available in the learning phase and therefore offered larger rewards during learning, this was no longer true during the test phase; participants therefore had no reason to expect previously rewarding groups to provide a larger reward in the test phase. A participants' tendency to choose based on prior generosity would therefore reflect a goal-directed process (because prior generosity is predictive of sharing), whereas their tendency to choose based on prior reward could be interpreted as reflecting a habit-like process (because, with the point pool known, prior reward is now irrelevant to predicted sharing). The persistence of group-based reward associations in guiding choice, even when such associations are no longer goal-relevant, would represent the hallmark of a habit-like behavior. Evidence for this persistence would therefore suggest that instrumental learning can give rise to relatively habitual tendencies of group-based interaction (Wood, 2017).

6.1. Method

6.1.1. Participants

Eight-two undergraduate students (76 female, 6 male) participated, with an additional eighteen excluded from analysis due to program failures during the experiment or failure to meet our inclusion criteria described previously. Sample size was determined by aiming to collect data from at least 80 participants, as in Study 2, with data collection continuing until the end of the semester.

6.1.2. Procedure

As in Studies 1 and 2, participants completed a learning phase and test phase of a "sharing game." In the learning phase, participants repeatedly selected a partner on each round. Unlike the previous experiments, however, feedback revealed two pieces of information: (a) how many points the player chose to share with them, as well as (b) the pool of points that player had available (Fig. 6b). Hence, this feedback simultaneously conveyed the absolute reward value of interacting with the player as well as their generosity. The average number of points shared and average proportion of points shared varied with university affiliation. Critically, these quantities were orthogonal across the groups, such that members of one group were rewarding but not generous, members of another group were generous but not rewarding, and so on. During the learning phase (162 trials), participants saw each possible pairing of groups an equal number of times. Three avatars (i.e., group members) were encountered from each group during learning.

The proportion shared by a target allowed participants to infer the target's generosity, whereas the amount shared by a target represents the amount of monetary reward. By manipulating the pools of money different targets had available, two targets could offer equivalent proportions but provide different levels of reward. In past work, reactions to reward feedback and generosity feedback have been found to rely on different neural pathways, with reward feedback involving neural regions linked to reward processing (e.g., ventral striatum) but generosity feedback further involving neural regions linked to social impression updating (Hackel et al., 2015). Moreover, people rely on generosity feedback more when learning about other humans as opposed to slot machines, further distinguishing these two types of feedback (Hackel et al., 2020).

In the subsequent test phase (180 trials), participants again made choices involving all possible pairings of groups, but with three changes from the learning phase (Fig. 6c). First, as in Studies 1 and 2, participants saw no further feedback; they were told they would find out how much they won at the end of the task. Second, participants again viewed both familiar and novel faces from each group (in separate pairs), allowing us to test yet again whether participants generalized each kind of learning to novel group members. Three new avatars were encountered from each group, in addition to the original avatars.

Finally, participants were told that, for the test phase, each avatar

had an equal number of points to share on each round (100 points). Critically, this last instruction rendered prior reward information irrelevant. For instance, groups B and C both shared 40% of the point pool on average during the learning phase, but group C typically had more points available than group B, allowing them to provide larger rewards. During the test phase, however, both groups had 100 points available on each trial, meaning that there was no longer any reason to prefer Group C; instead, a goal-directed learner should equally desire to interact with B and C. Indeed, prior work has found that the optimal strategy in the test phase is to ignore reward information and choose based only on generosity (Hackel et al., 2015). However, previously-formed reward associations might lead people to continue choosing Group C. As such, this design permitted us to test whether people continue to follow reward contingencies in a habit-like manner when choosing group members as interaction partners.

After the choice task, participants again rated the generosity of each avatar, completed feeling thermometer ratings toward each group, and reported whether they recalled seeing each avatar during the learning phase.

6.2. Results

6.2.1. Test phase choice

Did participants persist in choosing avatars in the test phase based on prior reward feedback, even when this feedback was statistically independent of generosity feedback and no longer earned them money? To address this question, we analyzed the likelihood of choosing an avatar (the avatar on the right side of the screen, selected arbitrarily), as a function of reward value and generosity, using mixed effects logistic regression. Predictors included the differences between the two groups shown on screen (right avatar – left avatar) in reward value and generosity, both of which were standardized within-participant to z-scores. We used this analysis strategy, rather than the analysis strategy used in Studies 1 and 2, because test phase choices in Study 3 could not be defined as simply "correct" or "incorrect;" participants could choose targets based on either reward or generosity. Instead, this analysis simply tests the extent to which participants used each form of feedback when making choices, as in prior work using this task (Hackel et al., 2015, 2020).

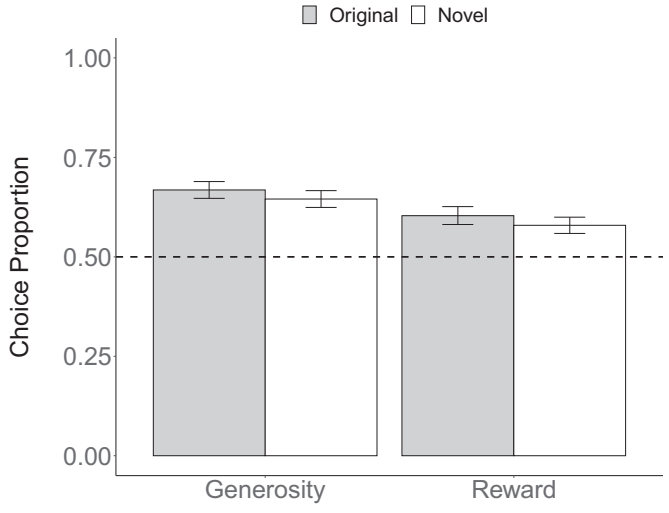

Fig. 7. Proportion of test phase choices in Study 3 for which participants selected the target onscreen that was higher in generosity and, independently, the proportion of choices for which participants selected the target that was higher in previous reward value, across original and novel group members. The dotted line indicates chance. Participants chose members of groups previously associated with reward for both original and novel faces. Error bars show standard error of the mean, with within-participants adjustment (Morey, 2008).

This analysis revealed main effects of reward value, $b = 0.45$, $SE = 0.15$, $z = 3.02$, $p = .003$, and generosity, $b = 0.96$, $SE = 0.16$, $z = 6.09$, $p < .001$, indicating that participants chose targets on the basis of both their reward and generosity (Fig. 7). That is, even though there was no longer any material benefit to choosing previously rewarding groups, participants continued to choose groups based on prior reward feedback in addition to prior generosity feedback. To test whether reward or generosity generalized to novel group members, we examined interactions of these factors with familiarity. This interaction was nonsignificant for both generosity, $b = -0.04$, $SE = 0.04$, $z = -1.24$, $p = .22$, and reward value, $b = -0.06$, $SE = 0.04$, $z = -1.55$, $p = .12$, indicating that novel group members were chosen similarly to original members of the same group. Indeed, the simple effect of reward value was positive for both familiar faces, $b = 0.48$, $SE = 0.15$, $z = 3.20$ $p = .001$, and novel faces, $b = 0.38$, $SE = 0.15$, $z = 2.54$, $p = .01$. Similarly, the simple effect of generosity was positive for both familiar faces, $b = 1.00$, $SE = 0.16$, $z = 6.31$, $p < .001$, and novel faces, $b = 0.88$, $SE = 0.16$, $z = 5.52$, $p < .001$. These findings provide evidence that people generalized prior reward feedback—in addition to trait feedback—to new group members, even when that reward feedback no longer signaled points earned. These results reveal that reward associations persisted in choice even when made irrelevant by changes in reward contingencies: Participants chose novel members of groups that previously provided large rewards, even though there was no reason to expect that these individuals would provide large rewards any longer.

6.2.2. Explicit impressions and attitudes

Post-task ratings of generosity were again analyzed by fitting a mixed effects linear regression predicting ratings for each avatar. Predictors included generosity (−1 = low, 1 = high) and reward value (−1 = low, 1 = high) of the avatar's group, as well as the familiarity (−1 = original, 1 = novel) of the avatar. This analysis revealed main effects of generosity, $b = 0.45$, $SE = 0.07$, $t(70) = 6.89$, $p < .001$, and reward value, $b = 0.28$, $SE = 0.08$, $t(70) = 3.66$, $p < .001$, indicating that explicit impressions of generosity were influenced by feedback about both generosity and reward value (Fig. 8a).

Participants applied impressions based on generosity feedback and reward feedback to original group members and generalized it to novel group members. Simple effects analysis revealed that reward feedback influenced impressions of original group members, $b = 0.33$, $SE = 0.08$, $t(91.15) = 4.05$, $p < .001$, and novel group members, $b = 0.23$, $SE = 0.08$, $t(91.15) = 2.78$, $p = .007$. The interaction of Reward and Familiarity was not significant, $b = -0.05$, SE = 0.03, $t(1348) = -1.77$, $p = .08$. Similarly, generosity feedback influenced impressions across original faces, $b = 0.53$, $SE = 0.07$, $t(99.99) = 7.42$, $p < .001$, and novel faces, $b = 0.36$, $SE = 0.07$, $t(99.99) = 5.03$, $p < .001$, although a Group x Familiarity interaction indicated a relatively larger influence on impressions of familiar faces, $b = -0.09$, $SE = 0.03$, $t(70) = -2.78$, $p = .007$. Together, these results show that participants formed positive trait impressions of groups not only based on feedback about the generosity they displayed but also based on feedback about the rewards they provided, and these impressions extended to novel group members.

To verify that participants formed overall attitudes toward each group based on instrumental learning, we again examined feeling thermometer ratings. Participants expressed more positive attitudes toward groups whose members were more generous, $b = 11.83$, $SE = 1.21$, $t(70) = 9.74$, $p < .001$, and more rewarding, $b = 4.68$, $SE = 1.61$, $t(70) = 2.91$, $p = .00$ (Fig. 8b). Thus, participants' attitudes toward social groups depended on group members' earlier generosity as well as rewards.

Finally, to determine whether reward feedback influenced choices in a manner distinct from its influence on explicit judgments, we again tested whether attitudes and impressions related to choices in the test phase. We added the difference in feeling thermometer scores for each group on screen (right group - left group) as a predictor of test phase choice, as well as the difference in mean generosity ratings given to each

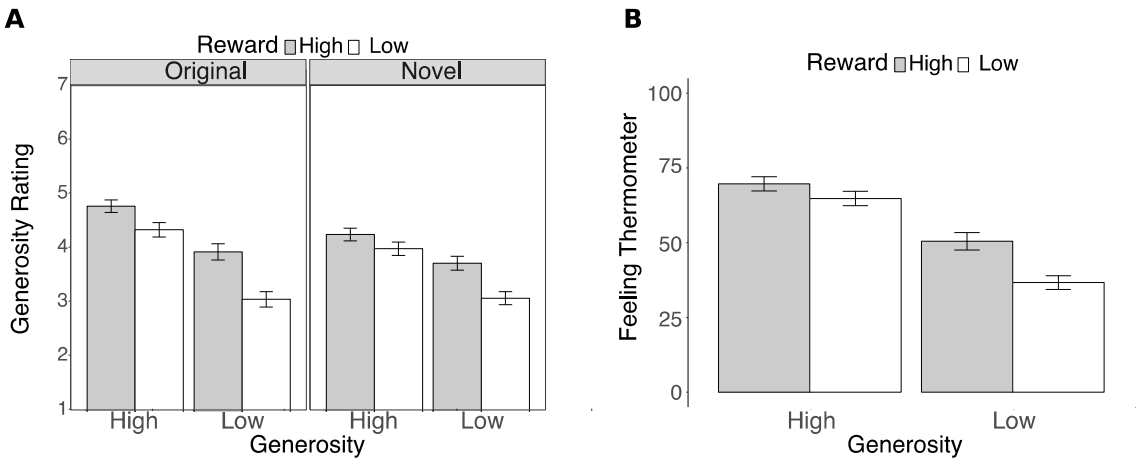

Fig. 8. Explicit attitude and impression ratings in Study 3. (a) Participants rated individual group members as more generous if their group had been associated with higher generosity feedback and if their groups had been associated with higher reward feedback, across original and novel exemplars.
(b) Participants had more positive attitudes toward groups that were associated with higher reward and groups that were associated with higher generosity.

group. Attitudes strongly predicted choices, $b = 1.05$, $SE = 0.18$, $z = 6.09$, $p < .001$, as did impressions of generosity, $b = 0.44$, SE = 0.08, $z = 5.80$, $p < .001$. The effect of manipulated generosity was no longer significant, $b = 0.09$, $SE = 0.10$, $z = 0.86$, $p = .39$, suggesting that the effect of generosity feedback on choice strongly overlapped with its effect on explicit ratings. In contrast, however, a small effect of reward feedback on choices remained when adjusting for explicit ratings, $b = 0.15$, $SE = 0.07$, $z = 2.01$, $p = .04$. These findings suggest that reward feedback may have shaped choices in a manner not fully overlapping with its impact on explicit attitudes or impressions, consistent with Studies 1 and 2 and the possibility of an implicit influence.

6.3. Discussion

Study 3 was designed to demonstrate the role of instrumental reward learning in the formation of group choice tendencies, independent of trait inferences that may simultaneously be formed during interactions with individual group members. To this end, it did so by using a task that experimentally dissociated reward feedback from trait feedback: Participants learned about groups that varied independently in trait-level generosity and material reward value. We found that participants generalized learning about reward value to novel group members, consistent with an instrumental learning mechanism based on experiences of reward, in addition to generalizing learning about generosity. That is, participants' tendency to choose novel partners was influenced not only by a group's generosity but also by the reward value of group members in prior interactions. These findings provide additional, and more direct, support for our hypothesized role of instrumental reward learning by demonstrating that reward feedback influences choices even when it is experimentally isolated from trait feedback.

Next, this study provides initial evidence for a habit-like effect of reward learning in intergroup interactions. Reward contingencies changed in the test phase, such that prior reward learning was rendered irrelevant to participants' goals. Nonetheless, participants continued to choose members of previously rewarding groups even though there was no longer any financial incentive to do so. This finding is consistent with the proposal that reward associations persist in social choices in a manner that may include the contribution of habits (Amodio, 2019; Amodio & Ratner, 2011; Hackel et al., 2019). Altogether, Study 3 revealed that reward learning has a unique impact on choices and attitudes toward social groups, relative to learning about the generosity others display, and that these reward associations persist in choice in a potentially habit-like manner.

7. Study 4

Study 4 was designed to address two final goals. First, we aimed to assess group-based learning in the presence of individuating information about group members. In Studies 1–3, all members of a group shared at the same rate, offering no individuating information for participants to use. In contrast, Study 4 participants learned about individuals who varied around a group mean. This design allowed us to test whether participants learned about individuals *and* groups when both types of information were present.

Second, we developed a new cover story to further explain the contingency change in the test phase, in order to rule out any concern that participants persisted in choosing rewarding targets because they were confused about the task. This cover story framed the task in terms of competence, rather than generosity: participants learned about "workers" who varied in their competence in answering trivia questions as well as in the reward feedback they provided. To ensure clarity, we (a) developed an intuitive explanation for why point pools varied across players and why contingencies changed in the test phase, (b) required participants to correctly answer an attention check question before the test phase about the change in contingencies, and (c) added a comprehension check after the task to ensure that participants understood the contingency change. These efforts allowed us to test persistence of group-based reward associations in choice while reducing the chances of participant mis-comprehension. Study 4 was pre-registered on AsPredicted.org (https://aspredicted.org/PS4_45Z).

7.1. Method

7.1.1. Participants

One hundred forty participants (63 female, 73 male, 2 non-binary, 2 no gender response) were recruited using the online platform Cloud Research. However, due to a computer error, the specific face stimuli shown during the test phase did not record properly for two participants, who were excluded from analyses of individuated target choice in the test phase data but not other analyses. In addition, post-task rating data did not properly record for 15 participants, who were excluded from analyses of post-task ratings but not other analyses. Sample size was determined by aiming to collect data from at least 112 participants, based on a simulation-based power analysis for the simple effect of reward among new faces in the Study 3 test phase. Additional participants were recruited to account for potential exclusions.

7.1.2. Procedure

The procedure was identical to that of Study 3— a learning phase, a

test phase, and post-task ratings of targets and groups—with three exceptions. First, the cover story framed the task in terms of competence rather than generosity, and, second, it offered a more intuitive explanation for reward feedback varying across players. Participants were told that previous MTurk workers had been assigned to teams for a trivia competition. As part of this competition, they had to answer questions such as "How many phases does the moon have?" After answering the question, these players had to choose a square from a scratch-off "lottery ticket," which revealed how many points were at stake for that round. Based on how close they were to the answer, they would earn a proportion of the point total on the lottery scratch-off. Participants were given an example of this process (see Supplemental Materials for instructions and sample images shown to participants). Next, participants were told that they would choose one player on each round and earn however many points that player had earned for their team. The instructions emphasized that lottery tickets could have high or low values, and as a result, participant bonuses would depend on both the number of points the chosen player had at stake from the scratch-off lottery and the proportion of those points earned through the trivia game.

Before completing the test phase, participants were told that they were now entering a "double jeopardy round." In this round, all scratch-offs were worth 1000 points. As in Study 3, this instruction changes the reward contingencies in the test phase: players who previously received larger amounts in the scratch-off would now have the same amount as everyone else. There would therefore be no reason to choose players based on their previous reward value; if two players have 1000 points at stake, then the one who tends to answer trivia questions more competently would provide a larger reward.

We instantiated an attention check to ensure participants comprehended this contingency change before they could progress to the test phase. Specifically, participants had to correctly answer a question indicating that each target would have 1000 points available on each round; if they did not answer correctly, they were informed of the correct answer and had to answer again. Finally, at the end of the task, participants answered a comprehension question that tested whether they understood that players who had larger point pools during the learning phase would have the same number as other players during the test phase. This second question did not require a correct response to continue; instead, it simply measured participant comprehension after the task.

In addition to changing the cover story and requiring comprehension checks, Study 4 introduced variability within groups: during the learning phase, individual group members varied around a group mean. Analogous to Study 3, teams varied independently in the average reward they provided (40 or 80 points) and the average competence they displayed (earning 0.40 or 0.80 of the point total). Participants learned about two players from each team during the learning phase. Unlike Study 3, however, one player from each team earned a reward 5 points higher and a proportion 0.05 higher than the group mean, and the other player earned a reward 5 points lower and a proportion 0.05 lower than the group mean. Therefore, one player was more worthwhile than the other along both dimensions in each team. This change allowed us to test participant learning about groups in the face of differences between members. Participants completed 96 trials of the learning task described in Study 3, in which they learned about two members from each of four teams.

After the test phase, participants again completed ratings of explicit impressions of players and attitudes toward groups. Unlike Study 3, participants rated impressions of both traits and rewards. Specifically, participants were asked to rate how competent each player was at the trivia task and how good each player was at getting large point values in the scratch-off lotteries. This question was added to once again test whether participants understood that these variables were distinct.

7.2. Results

7.2.1. Task comprehension

We first ensured that participants understood the change in task contingencies in the test phase. Specifically, we examined whether they correctly indicated that players with many "scratch-off" points in the learning phase had the same amount as other players in the test phase. Of the 125 participants for whom we had data for this question, 89% gave the correct answer, demonstrating high comprehension. Among the 14 participants with incorrect responses, only four indicated that some players had larger scratch-off pools at both learning and test; seven participants indicated that all players had the same scratch-off amounts during both learning and test, and three indicated that all players had equal amounts at learning but some had more at test. These data indicate that participants overwhelmingly held the correct model of the task structure, and of the small minority who did not, half held a task model that still would not predict some players having larger pools at test. An exploratory analysis including only participants who responded correctly to the comprehension check did not change inferences (Table S1).

7.2.2. Test phase choice

We first asked whether participants successfully learned about individual variability within groups. We analyzed the likelihood of choosing original group members during the test phase. (This analysis was restricted to original group members because participants did not learn about individual variability for novel members). We used the same mixed effects regression approach described in Study 3, analyzing choice as a function of reward value and competence. However, in addition to including regressors describing the average reward value and average competence of each player's group, we included a regressor accounting for individual deviations from those group means. Avatars were scored on whether they deviated positively or negatively from their group mean (1 or − 1). The difference between avatars on this score (right - left side of the screen) served as a regressor in the analysis of choice. This analysis revealed that participants were more likely to choose targets who deviated positively, rather than negatively, from their group means, $b = 0.21$, $SE = 0.04$, $z = 5.01$, $p < .001$. Participants thus learned individuating information about original group members, allowing us to assess group-based choice in the presence of individual-based learning.

We therefore next asked whether participants persisted in choosing avatars from rewarding groups in the test phase, even when the task clearly indicated this feedback was statistically independent of competence feedback and would not carry over from learning. We used the same mixed effects regression strategy described in Study 3, predicting choice as a function of competence, reward, and avatar novelty.

This analysis revealed main effects of reward value, $b = 0.97$, $SE = 0.12$, $z = 8.14$, $p < .001$, and competence, $b = 1.58$, $SE = 0.14$, $z = 11.34$, $p < .001$ (Fig. 9). That is, even though the task explained that there was no longer any material benefit to choosing previously rewarding groups, participants continued to choose groups based on prior reward feedback in addition to prior competence feedback. Unlike Study 3, we observed a small but significant interaction of avatar familiarity with competence, $b = -0.10$, $SE = 0.03$, $z = -3.00$, $p = .003$. We did not observe a significant interaction of avatar familiarity with reward value, $b = -0.05$, $SE = 0.03$, $z = -1.84$, $p = .07$. Simple effects analysis revealed that the effects of competence and reward were significant across original faces (Competence: $b = 1.68$, $SE = 0.14$, $z = 11.70$, $p < .001$; Reward: $b = 1.02$, $SE = 0.12$, $z = 8.33$, $p < .001$) and novel faces (Competence: $b = 1.48$, $SE = 0.14$, $z = 10.38$, $p < .001$; Reward: $b = 0.92$, $SE = 0.12$, $z = 7.52$, $p < .001$). These findings replicate those of Study 3: participants generalized prior reward feedback—in addition to trait feedback—to new group members, and they did so in a persistent manner despite having no reason to expect that these individuals would provide larger rewards than others.

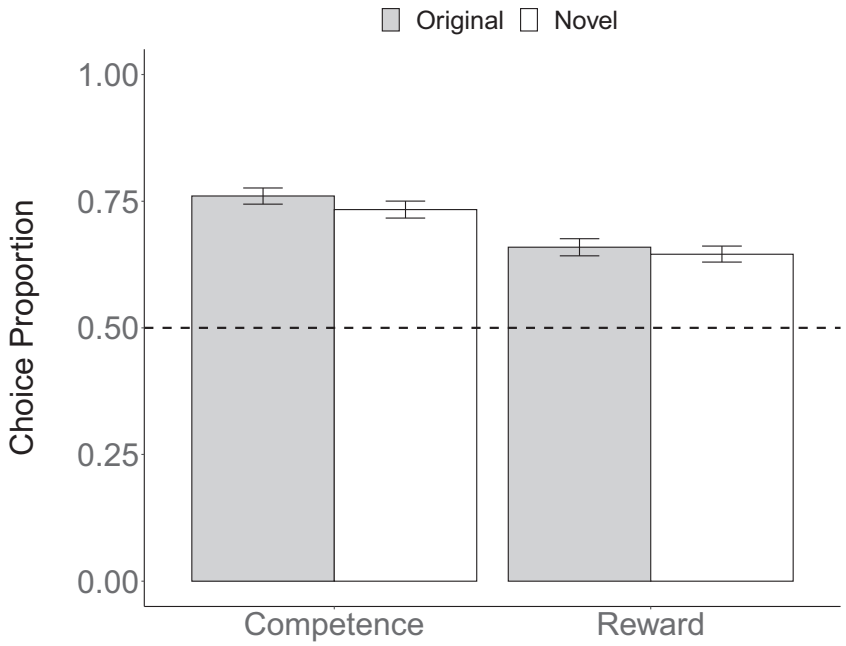

Fig. 9. Proportion of test phase choices in Study 4 for which participants selected the target onscreen that was higher in competence and, independently, the proportion of choices for which participants selected the target that was higher in previous reward value, across original and novel group members. The dotted line indicates chance. Participants chose members of groups previously associated with reward for both original and novel faces. Error bars show standard error of the mean, with within-participants adjustment (Morey, 2008).

7.2.3. Explicit impressions and attitudes

Post-task explicit impressions were again analyzed by fitting a mixed effects linear regression predicting ratings for each avatar. These ratings included impressions of the avatar's competence and of the avatar's success in gaining many points from the scratch-off lotteries. Predictors included competence (−1 = low, 1 = high) and reward value (−1 = low, 1 = high) of the avatar's group, familiarity of the avatar (−1 = original, 1 = novel), and rating type (−1 = points, 1 = competence).

This analysis revealed a strong Competence x Rating Type interaction, $b = 0.22$, $SE = 0.04$, $t(12.92) = 5.84$, $p < .001$, indicating that participants relied on competence feedback mainly when judging trivia competence (simple effect: $b = 0.65$, $SE = 0.06$, $t(237.13) = 10.71$, $p < .001$), and far less so when judging scratch-off lottery success (simple effect: $b = 0.20$, $SE = 0.06$, $t(237.25) = 3.38$, $p = .0008$). Participants therefore understood that trivia competence was distinct from scratch-off success. We did not observe a significant Reward x Rating Type interaction, $b = -0.05$, $SE = 0.02$, $t(124.03) = -1.91$, $p = .06$. Simple effects analysis revealed that reward feedback influenced impressions of both trivia competence, $b = 0.30$, $SE = 0.05$, $t(199.54) = 6.08$, $p < .001$, and scratch-off success, $b = 0.39$, SE = 0.05, $t(199.70) = 8.01$, $p < .001$, to similar degrees. Thus, reward feedback carried affective weight that influenced impressions of targets. Altogether, although participants broadly understood the distinction between the two types of feedback in the task, they also showed a degree of spillover, such that reward feedback influenced impressions of competence in addition to impressions of scratch-off success.

Simple effects analysis further revealed that, in ratings of trivia competence, reward feedback did not have a significant interaction with avatar familiarity, $b = -0.05$, $SE = 0.02$, $t(358.99) = -1.97$, $p = .05$. That is, reward feedback influenced judgments of both original avatars, $b = 0.35$, $SE = 0.05$, $t(293.03) = 6.34$, $p < .001$, and new avatars, $b = 0.25$, $SE = 0.05$, $t(292.63) = 4.64$, $p < .001$. Competence feedback had a stronger interaction with avatar familiarity, $b = -0.13$, $SE = 0.03$, $t(258.48) = -4.73$, $p < .001$: competence had a stronger effect on judgments of original avatars, $b = 0.78$, $SE = 0.07$, $t(332.65) = 11.70$, $p < .001$, than new targets, $b = 0.52$, $SE = 0.07$, $t(332.29) = 7.79$, $p <$

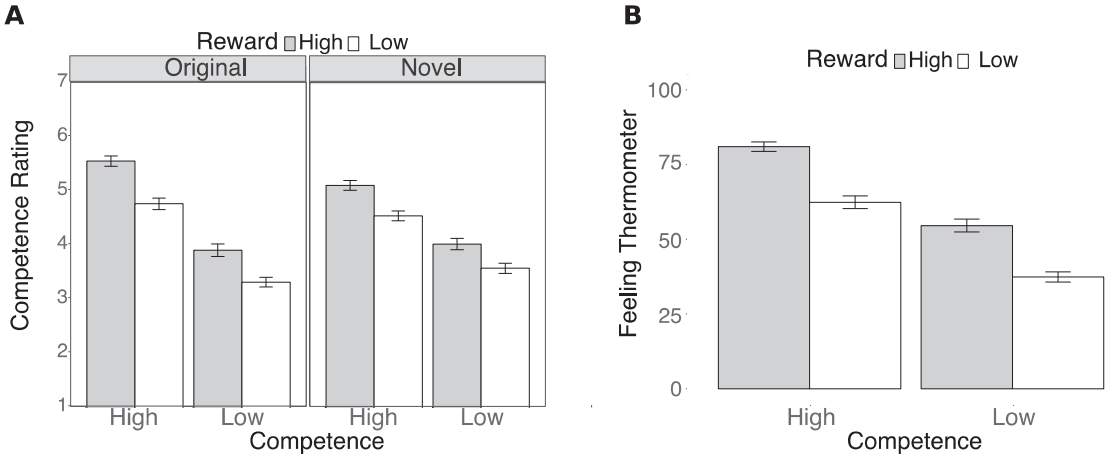

Fig. 10. Explicit attitude and impression ratings in Study 4. (a) Participants rated individual group members as more competent if their group had been associated with higher competence feedback and if their groups had been associated with higher reward feedback, across original and novel exemplars. (b) Participants had more positive attitudes toward groups that were associated with higher reward and groups that were associated with higher competence.

.001. Thus, participants applied impressions based on reward feedback to new and original group members.

Next, to verify that people formed overall attitudes toward each group based on instrumental learning, we again examined feeling thermometer ratings. Participants expressed more positive attitudes toward groups whose members were more competent, $b = 12.84$, $SE = 1.16$, $t(135) = 11.05$, $p < .001$, and more rewarding, $b = 8.94$, $SE = 1.06$, $t(135) = 8.47$, $p < .001$ (Fig. 10). Thus, participants' attitudes toward social groups depended on group members' earlier generosity as well as rewards.

Finally, to explore whether reward feedback influenced choices in a manner distinct from its influence on explicit judgments, we again tested separately whether attitudes and impressions related to choices in the test phase. We added the difference between groups onscreen in (a) generosity ratings, (b) ratings of lottery success, and (c) feeling thermometer ratings to our regression model predicting choice. Choices were influenced by attitudes, $b = 1.73$, $SE = 0.18$, $z = 9.66$, $p < .001$, and explicit impressions of trivia competence, $b = 0.81$, $SE = 0.13$, $z = 6.21$, $p < .001$. Nonetheless, we continued to observe effects of manipulated competence, $b = 0.81$, $SE = 0.13$, $z = 6.21$, $p < .001$, and manipulated reward value, $b = 0.20$, $SE = 0.07$, $z = 2.94$, $p = .003$. These findings again suggest that reward feedback shaped choices in a manner not fully overlapping with its impact on explicit attitudes or impressions.

7.3. Discussion

Study 4 was designed to demonstrate instrumental reward learning in a second trait domain—competence—while introducing individual variability within groups. To this end, it employed task instructions that described the proportion won by players as reflecting competence in a trivia task and the reward won as reflecting the amount of points available from a scratch-off lottery. Although participants learned which individuals were better or worse within groups, participants also learned about the competence and reward value of groups overall and generalized this learning to novel group members.

In addition, Study 4 was designed to provide stronger evidence of persistent, habit-like tendencies. To do so, task instructions offered a more intuitive explanation of why some players offered larger rewards than others during learning but not at test. Specifically, the points at stake for each question during learning were determined by a scratch-off lottery that could give a small or large amount, whereas in the test phase, every square on the scratch-off ticket would give 1000 points. This design was explained at length, and participants understood this change, as revealed in a comprehension check question. Crucially, this change meant that only a player's competence should predict the reward acquired from choosing them during the test phase. Nonetheless, participants again continued to choose members of previously rewarding groups, even though there was no longer any financial incentive to do so. Altogether, the results of Study 4 demonstrate instrumental learning about groups in a new domain (competence) in the presence of individuating information about group members, and they provide stronger evidence that these reward associations persist in choice in a habit-like manner.

8. General discussion

Across four studies, instrumental learning in direct interactions with individual group members formed the basis for group-based interaction tendencies. That is, participants' rewarding experiences with individuals influenced the reward value associated with those individuals' groups, and this value was reflected in group attitudes, impressions of generosity, and choices to interact with or avoid novel members of the same groups.

Our finding of an instrumental basis for a group attitude suggests a mode of prejudice formation distinct from prior conceptualizations rooted in passive forms of learning such as instruction, observation, or evaluative conditioning. In each study, participants learned the value of different groups via rewarding interactions with individuals. Thus, rather than passively witnessing or being told about the character traits of others, participants learned about the value of action involving each group—a key feature distinguishing instrumental learning from other modes of learning. Although people do often form prejudices in the absence of direct interaction, theory and evidence in cognitive neuroscience give reason to think that interactive learning stems from distinct mechanisms and carries distinct consequences relative to more passive forms of learning (Amodio, 2019; Foerde et al., 2006; Hackel et al., 2019; Poldrack et al., 2001; Wood, 2017). The present work identifies this reward-based perspective as another route to intergroup prejudice, complementing other forms of prejudice through cultural transmission studied in prior work.

This group-based value was generalized to novel group members with whom participants had never interacted—a hallmark of prejudice—such that participants chose novel members of social groups based on past rewarding feedback in individual interactions with other group members. Moreover, in Study 2, we found that this generalization occurred in approach learning and avoidance learning, which have been linked to dissociable neural substrates (Frank et al., 2004, 2007). That is, participants learned to approach groups associated with high value and to avoid groups associated with low value, demonstrating that both kinds of instrumental associations can be applied to social groups.

Finally, we found that reward feedback was generalized to novel group members even when manipulated independently of trait feedback. In Studies 3 and 4, the reward a group provided (the amount of money provided) was experimentally dissociated from the generosity or competence a group displayed (the proportion of money shared or earned), thus ensuring that the manipulation of reward feedback was not confounded with trait information. Results revealed that participants chose to interact with novel members of groups that had been previously rewarding, independent of their previous generosity or competence. Thus, both when measuring participants' impressions and attitudes (Studies 1–3) and when experimentally controlling for trait feedback (Studies 3–4), we found that reward feedback shaped participants' decisions to interact with novel group members. These findings thus identify a reward-learning pathway that gives rise to intergroup behavior and expand the role for learning processes in the formation of group value through interaction.

Importantly, interactions with group members can be more or less rewarding even in the absence of real group differences in behavior. For instance, when policies promote inequality, some groups end up with greater material and social capital than others, allowing them to offer more rewarding interactions despite identical character or ability. Studies 3 and 4 mirror this type of inequality, in which two individuals have equivalent generosity or competence but one of them won more money in a lottery and offers more rewarding interactions. Analogously, Hackel and Zaki (2018) found that people reciprocate more with wealthy individuals, propagating inequality between individuals through reward learning; the present findings suggest the same may be true of inequality between groups. Second, due to de-facto segregation, people often have little experience with other groups and experience anxiety in intergroup interactions. This anxiety may serve as a negative reinforcer of interaction (Wood, 2017). Finally, in Study 4, participants formed group-based value representations from interactions with only two individuals per group. Given extensive segregation between groups in many locations throughout the world (Enos, 2017), people may often interact with few members of other groups—and if these interactions happen to be negative, then the interactions may be sufficient to produce a negative value representation of the group. Moreover, if this negative value leads people to avoid the group, then people will not learn to correct this value representation through positive experiences (Allidina & Cunningham, 2021; Fazio, Eiser, & Shook, 2004). Altogether, an implication of the present work is that prejudice can arise from experiences of reward even if reward is incidental to a group's

character—for instance, rooted in societal inequality, segregation, intergroup anxiety, or unrepresentative interactions with particular individuals.

8.1. Reward learning in complex social contexts

In the present work, participants learned about novel groups, allowing us to isolate the effects of group categorization in the absence of pre-existing stereotypes, conflict, or prejudice (Tajfel, Billig, Bundy, & Flament, 1971). Yet, these other intergroup processes likely interface with reward learning. First, as noted above, people often feel anxiety during interactions when groups have a history of conflict, and this anxiety may itself serve as a negative reinforcer (Wood, 2017). Second, pre-existing stereotypes may alter reward learning about group members, biasing learning from individual interactions (Stillerman et al., 2020). Finally, once strong group representations have been formed, individuals who deviate from that group value may be sub-typed as part of a sub-category (Hewstone, Macrae, Griffiths, Milne, & Brown, 1994). Computationally, this process may relate to models of social structure learning (Gershman & Cikara, 2020; Lau, Pouncy, Gershman, & Cikara, 2018), wherein people assign individuals into latent groups based on prior expectations about groups and individual behavior. Understanding how pre-existing knowledge enhances or diminishes generalization of reward learning offers a fruitful direction for future research. Nonetheless, the present findings identify reward learning as another process influencing intergroup behavior and attitudes beyond conceptual forms of learning.

Individuals in the present research were also categorized as part of one group, but people belong to many social groups at once. At different times, different group categorizations will be salient to perceivers due to chronic accessible or situational cues (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987); the same individual may sometimes be categorized as a woman, a professor, or a European. The way in which people perceive and evaluate others depends on which categories are salient at a given moment (Hackel, Coppin, Wohl, & Van Bavel, 2018; Hornsey, 2008; Turner et al., 1987; Van Bavel & Cunningham, 2009). As a result, changing which social categories are salient may alter group-based value representations, suggesting a potential source of flexibility in group-based value. This possibility presents a fruitful avenue for future research.

Finally, in the present research, participants made decisions about novel group members without obtaining new, individuating information about those group members, allowing a test of pure group-based approach and avoidance. Future research can test whether group-based reward associations persist in the face of individuating information about novel group members, as has been observed in the context of implicit attitudes (McConnell, Rydell, Strain, & Mackie, 2008).

8.2. A Role for habits in intergroup relations

Discrimination has been likened to a habit, in the sense that people often act in discriminatory ways despite egalitarian goals (Devine, 1989). Yet, little research has tested whether intergroup behavior involves instrumental learning processes that give rise to habits. Habits refer to associations between a context and a response, which can be cued and enacted even in the absence of intention (Verplanken & Orbell, 2021; Wood & Rünger, 2016). As a result, habits could underlie discrimination in intergroup interactions if people repeat particular responses with social groups (e.g., approaching some, avoiding others) that no longer match their intentions. Indeed, Wood (2017) theorized that people may form habits in intergroup settings by developing habitual responses to interact with other groups as they experience positive and negative rewards during cross-group interactions, whether these involve material outcomes (e.g., receiving gifts) or social ones (e.g., experiencing anxiety). In turn, people easily perceive and categorize others' features reflecting group membership, which may trigger these relatively automatic responses in novel interactions. To the extent this process is involved, interventions for prejudice might be better suited trying to change behavior and experience than attitudes and impressions, given that habitual behaviors depend on contextual cues in environments rather than goals or intentions (Neal, Wood, Wu, & Kurlander, 2011).

This possibility that people could form intergroup habits served as a motivating consideration in studying instrumental learning—the key learning process by which habits are formed. By linking group-based interaction choices to instrumental learning, our findings provide initial evidence that could support such habit-like tendencies in cross-group interaction. Unlike more passive forms of learning, instrumental learning can give rise to habits, wherein people persist in previously rewarded behaviors even when such behaviors will no longer attain desired rewards. For instance, during social interactions, people form model-free reward associations that guide their social choices and attitudes —a form of learning that can give rise to persistent, habit-like patterns of choice (Hackel et al., 2019) and to implicit attitudes toward social groups (Kurdi, Gershman, & Banaji, 2019). Alternatively, reward feedback can prompt people to repeat actions within a given context, and this mere repetition might promote habit formation (Miller et al., 2019). Through either pathway, people could form habits to approach or avoid social groups through generalizing instrumental learning—a proposition that depends on first identifying instrumental learning about groups. By characterizing instrumental learning about groups, the present work thus supports the possibility of intergroup habits.

Although our studies were not designed to test the role of habit directly, they provide suggestive evidence. First, although the task used in Studies 1 and 2 does not directly dissociate habit and goal-directed choice, behavior in this task is thought to primarily reflect instrumental associations acquired by the striatum during instrumental learning (Doll et al., 2016; Frank et al., 2007; Jocham et al., 2011). This form of learning may support the formation of habitual responses (Foerde et al., 2006). Second, in each of the present studies, reward feedback influenced behavior even when statistically accounting for participants' explicit attitudes and impressions. Although this analysis cannot rule out the influence of other unmeasured variables, it is consistent with an implicit (i.e., direct) influence on behavior that may reflect habit. Third, in Studies 3–4, participants persisted in choosing members of previously-rewarding groups even after reward contingencies changed, such that it was no longer beneficial to choose them. This finding resembles tests of contingency degradation—a classic marker of habits wherein animals continue to perform previously rewarded behaviors even after contingencies shift such that it is no longer rewarding to do so (Wood & Rünger, 2016). In Study 4, in particular, participants persisted in choosing members of previously rewarding groups even when they understood that these groups would no longer have any advantage in providing high levels of reward. These pieces of evidence suggest that the exploration of other hallmarks of habits (e.g., reward devaluation) in intergroup interactions offers a promising avenue for future work. Altogether, by demonstrating that instrumental learning promotes group-based attitudes and choice, our findings take a first step toward bridging intergroup relations and habitual learning processes, suggesting that a "discriminatory habit" could be more than a figure of speech.

This proposal offers new insights into interventions to decrease discrimination. Bias interventions often focus on changing people's motivations, intentions, beliefs, or attitudes. Although such changes can shape deliberate behaviors, they offer less effective routes to changing habits. Instead, interventions to change discrimination could draw on principles of habit formation. For instance, interventions could focus on disrupting contextual cues or situational affordances that trigger bias (Amodio & Swencionis, 2018; Wood, 2017; Wood & Neal, 2007). Alternatively, interventions could focus on creating "environmental friction" that makes biased behaviors more difficult to perform or

creating incentives that make egalitarian behaviors easy to perform (Wood & Neal, 2016). Devine and colleagues have proposed and tested an extensive "habit-breaking" intervention to reduce race- and gender-based prejudices that involves training to recognize cues for potential bias and then act in unbiased ways (Carnes et al., 2015; Devine et al., 2017; Devine, Forscher, Austin, & Cox, 2012). Although Devine et al.'s (2012) intervention addresses a broad set of processes beyond those related to habit per se and does not include habit-specific assessments, the inclusion of procedures that train participants to link specific intergroup cues to specific non-biased actions suggests that it may indeed affect habits. Whereas the term "habit" has been used more colloquially in past intergroup research, our analysis suggests that interventions targeting a more precise mechanism of habits may complement and enhance existing strategies that focus on changing attitudes, beliefs, and intentions.

At the same time, an instrumental learning lens also suggests that fruitful interventions may target broad social environments rather than individual behavior, given that social environments can make some behaviors more rewarding than others. For instance, in Studies 3–4 of the present work, some groups were endowed with larger point pools and therefore offered more materially reward interactions; participants interacted more with these groups. This pattern mirrors societal inequalities between groups, which may similarly set the stage for more or less rewarding interactions with members of different groups, leading people to associate different groups with high or low value. This view suggests a novel route by which societal inequalities promote individual bias, rooted in how individuals interact with their environments (Fiedler & Wänke, 2009). This type of bias may be particularly difficult to change, given that they may prompt habits and that people may not always realize their experiences reflect structural inequalities. As a result, an instrumental learning lens highlights broad societal inequalities as a target of intervention for intergroup bias, given that these inequalities constrain the playing field in which individuals learn through action and reward.

9. Conclusions

The present work identifies an instrumental learning mechanism that gives rise to attitudes toward social groups: through interactions with individuals, people form reward associations with social groups as a whole. This reward-based learning influences people's social choices, attitudes, and impressions, leading them to prefer to interact with new members of groups associated with previous rewarding experiences. This finding highlights a pathway by which prejudices can form and potentially change, rooted in active learning and choice rather than passive conceptual association.

More broadly, our findings suggest a role for multiple learning processes in social cognition (Amodio, 2019), including instrumental learning processes that give rise to social choices and attitudes (Hackel et al., 2019). As such learning experiences are repeated, people might form interaction habits that are enacted automatically with limited thought (Wood, 2017), thereby perpetuating group prejudices and societal inequalities even when people wish to act in less biased ways. These findings thus illuminate how social learning through reward feedback can support social behavior.

Open Practices

The data and analysis code for the current studies are available at https://osf.io/7nyaj/?view_only=d04aef5500c945b98eac1d7b5d10f04e. The preregistration for Study 4 is available at https://aspredicted.org/PS4_45Z

Acknowledgements

L.M.H. was supported by funding from the U.S. National Science Foundation (Award 1606959), W.W. was supported by a grant from the Templeton Foundation (Award 52316), and D.M.A was supported by grants from the U.S. National Science Foundation (Award 1551826) and the Netherlands Organisation for Scientific Research (VICI 016.185.058).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2021.104267.

References

Allidina, S., & Cunningham, W. A. (2021). Avoidance begets avoidance: A computational account of negative stereotype persistence. *Journal of Experimental Psychology: General*.

Amodio, D. M. (2019). Social Cognition 2.0: An interactive memory systems account. *Trends in Cognitive Sciences, 23*(1), 21–33.

Amodio, D. M., & Ratner, K. G. (2011). A memory systems model of implicit social cognition. *Current Directions in Psychological Science, 20*(3), 143–148.

Amodio, D. M., & Swencionis, J. K. (2018). Proactive control of implicit bias: A theoretical model and implications for behavior change. *Journal of Personality and Social Psychology, 115*(2), 255.

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology, 37*(4–5), 407–419.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Carnes, M., Devine, P. G., Manwell, L. B., Byars-Winston, A., Fine, E., Ford, C. E., … Magua, W. (2015). Effect of an intervention to break the gender bias habit for faculty at one institution: A cluster randomized, controlled trial. *Academic Medicine : Journal of the Association of American Medical Colleges, 90*(2), 221.

Crawford, M. T., Sherman, S. J., & Hamilton, D. L. (2002). Perceived entitativity, stereotype formation, and the interchangeability of group members. *Journal of Personality and Social Psychology, 83*(5), 1076.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology, 44*(1), 20.

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.

De Houwer, J. (2006). Using the Implicit Association Test does not rule out an impact of conscious propositional knowledge on evaluative conditioning. *Learning and Motivation, 37*(2), 176–187.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*(6), 853.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*(1), 5.

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology, 48*(6), 1267–1278.

Devine, P. G., Forscher, P. S., Cox, W. T., Kaatz, A., Sheridan, J., & Carnes, M. (2017). A gender bias habit-breaking intervention led to increased hiring of female faculty in STEMM departments. *Journal of Experimental Social Psychology, 73*, 211–215.

Doll, B. B., Bath, K. G., Daw, N. D., & Frank, M. J. (2016). Variability in dopamine genes dissociates model-based and model-free reinforcement learning. *Journal of Neuroscience, 36*(4), 1211–1222.

Dunsmoor, J. E., & Murphy, G. L. (2014). Stimulus typicality determines how broadly fear is generalized. *Psychological Science, 25*(9), 1816–1821.

Enos, R. D. (2017). *The space between us: Social geography and politics*. Cambridge University Press.

Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology, 87*(3), 293.

Fiedler, K., & Wänke, M. (2009). The cognitive-ecological approach to rationality in social psychology. *Social Cognition, 27*(5), 699–732.

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences, 103*(31), 11778–11783.

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.

Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science, 306*(5703), 1940–1943.

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 37*(7), 1297–1310.
Gershman, S. J., & Cikara, M. (2020). Social-structure learning. *Current Directions in Psychological Science, 29*(5), 460–466.
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.
Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology, 90*(1), 1.
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.
Hackel, L. M., Coppin, G., Wohl, M. J., & Van Bavel, J. J. (2018). From groups to grits: Social identity shapes evaluations of food pleasantness. *Journal of Experimental Social Psychology, 74*, 270–280.
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.
Hackel, L. M., Looser, C. E., & Van Bavel, J. J. (2014). Group membership alters the threshold for mind perception: The role of social identity, collective identification, and intergroup threat. *Journal of Experimental Social Psychology, 52*, 15–23.
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.
Hackel, L. M., & Zaki, J. (2018). Propagation of economic inequality through reciprocity and reputation. *Psychological Science, 29*(4), 604–613.
Heider, F. (1958). *The psychology of interpersonal relations*. Wiley.
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.
Hewstone, M., Macrae, C. N., Griffiths, R., Milne, A. B., & Brown, R. (1994). Cognitive models of stereotype change:(5). Measurement, development, and consequences of subtyping. *Journal of Experimental Social Psychology, 30*(6), 505–526.
Hornsey, M. J. (2008). Social identity theory and self-categorization theory: A historical review. *Social and Personality Psychology Compass, 2*(1), 204–222.
Jocham, G., Klein, T. A., & Ullsperger, M. (2011). Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *Journal of Neuroscience, 31*(5), 1606–1613.
Jones, E. E. (1985). Major developments in social psychology during the past five decades. In (3rd ed.)*, Vol. 1. The handbook of social psychology* (pp. 47–107). Random House.
Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization. In *, Vol. 55. Advances in experimental social psychology* (pp. 1–80). Elsevier.
Kurdi, B., Gershman, S. J., & Banaji, M. R. (2019). Model-free and model-based learning processes in the updating of explicit and implicit evaluations. *Proceedings of the National Academy of Sciences, 116*(13), 6035–6044.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2.0-32*.
Lau, T., Pouncy, H. T., Gershman, S. J., & Cikara, M. (2018). Discovering social groups via latent structure learning. *Journal of Experimental Psychology: General, 147*(12), 1881.
Lott, A. J., & Lott, B. E. (1974). The role of reward in the formation of positive interpersonal attitudes. In *Foundations of interpersonal attraction* (pp. 171–192). Elsevier.
Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology, 69*(3), 397.
McConnell, A. R., Rydell, R. J., Strain, L. M., & Mackie, D. M. (2008). Forming implicit and explicit attitudes toward individuals: Social group association cues. *Journal of Personality and Social Psychology, 94*(5), 792.
Miller, K. J., Shenhav, A., & Ludvig, E. A. (2019). Habits without values. *Psychological Review, 126*(2), 292.
Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology, 4*, 61–64.
Neal, D. T., Wood, W., Wu, M., & Kurlander, D. (2011). The pull of the past: When do habits persist despite conflict with motives? *Personality and Social Psychology Bulletin, 37*(11), 1428–1437.
Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science, 12*(5), 413–417.
Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology, 38*(6), 922–934.
Poldrack, R. A., Clark, J., Paré-Blagoev, E. A., Shohamy, D., Moyano, J. C., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature, 414*(6863), 546–550.
R Development Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9*(7), 545–556.
Ratner, K. G., & Amodio, D. M. (2013). Seeing "us vs. Them": Minimal group effects on the neural encoding of faces. *Journal of Experimental Social Psychology, 49*(2), 298–301.
Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: a systems of reasoning analysis. *Journal of Personality and Social Psychology, 91*(6), 995.
Stillerman, B., Lindström, B., Schultner, D., Hackel, L., Hagen, D., Jostmann, N., & Amodio, D. (2020). *Internalization of societal stereotypes as individual prejudice*. doi.10.31234/osf.io/mwztc.
Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1*(2), 149–178.
Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.
Uleman, J. S., & Kressel, L. M. (2013). A brief history of theory and research on impression formation. In *Oxford library of psychology. The Oxford handbook of social cognition* (pp. 53–73). Oxford University Press.
Van Bavel, J. J., & Cunningham, W. A. (2009). Self-categorization with a novel mixed-race group moderates automatic social and racial biases. *Personality and Social Psychology Bulletin, 35*(3), 321–335.
Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: Evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience, 23*(11), 3343–3354.
Verplanken, B., & Orbell, S. (2021). Attitudes, habits, and behavior change. *Annual Review of Psychology, 73*.
Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology, 47*(2), 237.
Wood, W. (2017). Habit in personality and social psychology. *Personality and Social Psychology Review, 21*(4), 389–403.
Wood, W. (2019). *Good habits, bad habits: The science of making positive changes that stick. Pan Macmillan*.
Wood, W., & Neal, D. T. (2007). A new look at habits and the habit-goal interface. *Psychological Review, 114*(4), 843.
Wood, W., & Neal, D. T. (2016). Healthy through habit: Interventions for initiating & maintaining health behavior change. *Behavioral Science & Policy, 2*(1), 71–83.
Wood, W., & Rünger, D. (2016). Psychology of habit. *Annual Review of Psychology, 67*.