Running head: Group-based instrumental learning

Group value learned through interactions with members: A reinforcement learning

account

Leor M. Hackel*

University of Southern California

Drew Kogon*

University of Southern California

David M. Amodio

New York University, University of Amsterdam

Wendy Wood

University of Southern California

*These authors contributed equally to this work.

9,740 words (includes abstract and main text)

Please direct correspondence to: Leor M. Hackel Department of Psychology University of Southern California 3620 South McClintock Ave Los Angeles, CA 90089 <u>lhackel@usc.edu</u>

Author Note

De-identified data for the experiments reported in this manuscript have been made available at: <u>https://osf.io/7nyaj/?view_only=fa4aaf673ab14b7f9f0539316fb82fbe</u>.

Abstract

How do group-based interaction tendencies form through encounters with individual group members? In three 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. Moreover, group-based reward values continued to predict interactions with novel members after controlling for explicit attitudes and impressions, suggesting that instrumental learning contributes to an implicit form of group-based choice. Experiment 3 further demonstrated that group-based reward effects on interaction choices persisted even when group reward value was no longer predicted of 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.

Keywords: Prejudice, instrumental learning, reward, attitudes, habit

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

When we interact with another person, we form attitudes and interaction patterns based on feedback they provide in the social exchange (Hackel et al., 2015; Lott & Lott, 1974). For instance, if another person shares resources with us, then this rewarding experience may lead us to like them (Hackel et al., 2019), choose to interact with them again (Hackel et al., 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.

Learning Value through Interaction: The Role of Instrumental Learning

Instrumental learning through direct interaction occurs when people perform actions toward another person and experience rewarding feedback. This feedback contributes incrementally to the *value* they associate with that person, 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 inferences about another person's qualities (e.g., trait impressions), 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 may involve habit (Daw et al., 2011; Miller et al., 2019; Rangel et al., 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 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). Although traditional models of social cognition focus on inferences about a partner's character traits (Heider, 1958; Jones, 1985; Uleman & Kressel, 2013), 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 et al., 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.

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 cannot choose to 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. Indeed, 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 et al., 1995) and basic forms of social perception such as visual face encoding and individuation (Hackel et al., 2014; Kawakami et al., 2017; Ratner & Amodio, 2013; Van Bavel et al., 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 its broader category (Dunsmoor & Murphy, 2014). 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 towards 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 et al., 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 et al., 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 et al., 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 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. 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.

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). 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.

Study 1

In Study 1, participants completed an economic game that was adapted from prior reinforcement learning tasks (Frank et al., 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.

Method

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 et al., 2015; Hackel et al., 2020; Hackel & Zaki, 2018). These a priori rules identified six participants to exclude, leaving 45 participants in the analyses. De-identified data from each study have been made available at:

https://osf.io/7nyaj/?view_only=fa4aaf673ab14b7f9f0539316fb82fbe.

Procedure

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, 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 (Figure 1a). Upon viewing the avatars, participants selected which of these two students they wished to interact with (Figure 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. 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.



Time

Figure 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.

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 et al., 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.

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 towards

groups as a whole, participants rated their attitudes towards 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 results did not change when adjusting for explicit memory. Thus, this variable is not discussed further.

Results

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 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.

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 (Figure 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 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 grouplevel representation, which then guided choices regarding novel group members.



Figure 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).

Explicit Impressions and Attitudes

Next, to determine whether reward learning carried forward into explicit impressions of individual targets and attitudes towards each group, we examined participants' post-task ratings of each avatar's generosity using mixed effects linear 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, Maechler, Bolker, & Walker, 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(Figure 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 mixed effects linear 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 (Figure 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 selfreports? To address this question regarding the effect of explicit attitudes, we first refit our regression model predicting test phase choice while accounting for feeling thermometer ratings of each group. Specifically, we added as a predictor the difference in participants' attitudes towards the two groups onscreen (higher reward group minus lower reward group), along with the interaction of attitudes with face familiarity. Although explicit attitudes were a significant predictor of choices, b = 0.06, SE = 0.02, z = 2.57, p = .01, the intercept remained significantly positive, b = 0.27, SE = 0.04, z = 7.37, p < .001, indicating that explicit attitudes did not fully account for the impact of reward feedback on choices. Next, to examine the role of trait impressions, we refit the regression model while adding 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. Again, the intercept remained significantly positive, b = 0.15, SE =0.03, z = 4.57, p < .001, even though impressions also related to choice, b = 0.11, SE = 0.02, z =6.63, p < .001. These findings suggest that prior reward feedback influenced subsequent choices independent of either explicit attitudes or impressions.



Figure 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 towards groups as a whole for groups that were associated with higher previous reward.

Discussion

Study 1 revealed that people learn about social groups through generalization of rewardbased 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 toward groups as a whole: Participants felt warmer toward groups associated with greater reward. 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 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 is one of the first pieces of evidence, to our knowledge, 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. We evaluated the nature of this direct effect further in our third study.

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 of 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 finegrained 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 et al., 2011). As a result, performance on transfer pairings may provide an even stronger index of instrumental learning than in Study 1. **Method**

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.

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.

Results

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-.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 = .89, SE = 0.13, z = 7.06, p < .001, but critically, this was true for both original and novel members (Figure 4). Simple effects analysis revealed that participants relied on prior reward learning both when choosing original avatars, b = .98, SE = 0.13, z = 7.43, p < .001, and when choosing novel avatars, b = .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 = -.09, SE = .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 = .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.



Figure 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 "Approach A" and "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).

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 (Figure 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 (Figure 5b). This finding verifies Study 1 in that participants formed attitudes towards groups as a whole based on instrumental reward feedback from its individual members.

Once again, however, explicit attitudes did not fully account for patterns of reward-based choice. When adding feeling thermometer scores for each group as a predictor in the analysis of test phase choices on transfer trials, they significantly predicted choice, b = 0.43, SE = 0.08, z = 5.50, p < .001, but the intercept remained significantly positive, b = .88, SE = 0.13, z = 6.99, p < .001, indicating that self-reported attitudes did not fully account for choice behavior. Thus, reward feedback shaped choice in a manner dissociable from its effect on explicit attitudes and choice.

Similarly, we examined whether explicit impressions of generosity accounted for choice patterns by entering as a predictor the difference in average generosity ratings (computed separately for original and novel exemplars) toward each group. Although impressions were a significant predictor of choice, b = 0.76, SE = 0.04, z = 19.07, p < .001, the intercept remained significantly positive, b = 1.03, SE = 0.14, z = 7.39, p < .001, indicating that impressions also did not fully account for choice. Altogether, explicit attitudes and impressions predicted choices but did not account for the impact of reward feedback on choice.

23



Figure 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 towards groups as a whole for groups that were associated with higher previous reward.

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, replicating Study 1, further suggesting

an implicit influence of reward feedback on behavior. The nature of this direct effect was evaluated in our final study.

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. In Studies 1 and 2, participants' explicit attitudes and impressions did not account for the impact of reward feedback on choice, suggesting 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 represented 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 average, revealing high generosity, and some avatars shared a large number of points, revealing their reward value (Figure 6a).

This task further allowed us to explore the extent to which reward-based decisions were rooted in goal-directed or relatively habitual behavior. In test phase trials of this task, point pool information was provided to participants when choosing between players. When point pools are known, a player's prior generosity, but not their prior reward, is a valid predictor of their monetary sharing. As such, a participants' tendency to choose based on prior generosity would reflect a goal-directed process (because only prior generosity is predictive of sharing), whereas their tendency to choose based on prior reward could be interpreted as reflecting a relatively habit-like process (because, with the point pool known, prior reward is irrelevant to predicted sharing). Because the persistence of group-based reward associations in guiding choice, even when such associations are no longer goal-relevant, represents the hallmark of a habit-like behavior, evidence for this persistence would suggest that instrumental learning can give rise to group-based interaction habits (Wood, 2017).

Method

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.

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 (Figure 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.



Figure 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.

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 (Figure 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 towards each group, and reported whether they recalled seeing each avatar during the learning phase.

Results

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 point pool, 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).

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 (Figure 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

29

large rewards, even though there was no reason to expect that these individuals would provide large rewards any longer.



Figure 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).

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 (Figure 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 = .33, SE = 0.08, t(91.15) = 4.05, p < .001, and novel group members, b = .23, SE = 0.08, t(91.15) = 2.78, p = .007; a marginally significant Reward x Familiarity interaction suggested reward might have had a larger influence on impressions of familiar faces, b = -.05, SE = .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 = -.007. Together, these results show that people 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 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 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 (Figure 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 separately whether attitudes and impressions related to choices in the test phase. First, we added the difference in feeling thermometer scores for each group on screen (right group - left group) as a predictor of test phase choice. Attitudes strongly predicted choices, b = 1.35, SE = 0.18, z = 7.36, p < .001, and did so more strongly for original than novel group members, b = -.13, SE = .05, z = -2.63, p = .008. The effect of manipulated generosity was no longer significant, b = 0.11, SE = 0.11, z = 1.06, p = .29, suggesting that the effect of generosity feedback on choice strongly overlapped with its effect on explicit attitudes. In contrast, however, a smaller effect of reward feedback on choices was marginally significant when adjusting for feeling thermometer scores, b = 0.14, SE = 0.08, z = 1.82, p = .07. Similarly, when we refit the models adjusting for the difference in mean generosity ratings given to each group, generosity ratings predicted choices, b = 1.02, SE = 0.14, z = 7.11, p < .001. However, the effect of generosity feedback remained significant, b = 0.40, SE = 0.11, z = 3.70, p < .001, and the effect of reward feedback remained marginally significant, b = 0.19, SE = 0.10, z = 1.88, p = .06. 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.



Figure 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 towards groups that were associated with higher reward and groups that were associated with higher generosity.

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 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.

General Discussion

Across three 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 previously unexplored 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.

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 Study 3, the reward a group provided (the amount of money shared) was experimentally dissociated from the generosity a group displayed (the proportion of money shared), 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. Thus, both when measuring participants' impressions and attitudes (Studies 1-3) and when experimentally controlling for trait feedback (Study 3), we found that reward feedback shaped participants' decisions to interact with novel group members. Moreover, reward feedback shaped group decisions even when statistically accounting for participants' explicit attitudes and impressions, suggesting instrumental learning may also have an implicit effect on group-based responses (Amodio & Ratner, 2011). These findings are consistent with an instrumental learning account of prejudice, thus identifying a reward-learning pathway that gives rise to intergroup behavior and expanding the role for learning processes in the formation of group value through interaction.

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 evidence 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 (Wood & Neal, 2007; 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 et al., 2011).

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 towards social groups (Kurdi et al., 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.

Although our studies were not designed to test the role of habit directly, they provide three pieces of suggestive evidence. First, the task used in Studies 1 and 2 is thought to primarily reflect instrumental associations acquired by the striatum during instrumental learning which may support the formation of habitual responses (Doll et al., 2016). Second, reward feedback influenced behavior even when statistically accounting for participants' explicit attitudes and impressions. This finding is consistent with an implicit (i.e., indirect) influence on behavior that may reflect habit. Third, in Study 3, participants persisted in choosing members of previouslyrewarding 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). 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 begin to bridge intergroup relations and habitual learning processes, suggesting that a "discriminatory habit" could be more than a figure of speech.

This proposal also 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 instead 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., 2012, 2017). To date, it is the only intervention shown experimentally

to produce long-lasting reductions in prejudice (Forscher et al., 2017). 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, a habit 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 Study 3 of the present work, some groups were endowed with larger point pools and therefore offered more materially reward interactions; participants subsequently interacted more with these groups. This pattern may mirror 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. For instance, when policies lead to the aggregation of wealth for dominant racial groups, people may be likely to encounter those groups in conditions that allow for materially rewarding interactions. 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, a habit 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.

Conclusions

The present work introduces an instrumental learning mechanism that gives rise to attitudes towards 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 suggests a novel pathway by which prejudices can be formed and potentially changed, 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.

References

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

- Forscher, P. S., C Mitamura, Dix, E. L., Cox, W. T., & Devine, Patricia G. (2017). Breaking the prejudice habit: Mechanisms, timecourse, and longevity. *Journal of Experimental Social Psychology*, 72, 133–146.
- 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 metaanalysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, 37(7), 1297– 1310.
- 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., 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.

- 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 *The handbook of social psychology* (3rd ed, Vol. 1, 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 *Advances in experimental social psychology* (Vol. 55, 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). ImerTest: Tests for random and fixed effects for linear mixed effect models (Imer objects of Ime4 package). R package

version 2.0-32.

- Lindström, B., & Tobler, P. (2018). Incidental ostracism emerges from simple learning mechanisms. *Nature Human Behaviour*, *2*(6), 405–414.
- 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.
- 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). *Tutorials 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? Metaanalytic tests of three mediators. *European Journal of Social Psychology*, 38(6), 922– 934.
- 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.
- 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., 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.
- 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.