Transmission of societal stereotypes to individual-level prejudices through instrumental learning

B. S. Stillerman[1], B. R. Lindström[2], D. T. Schultner[3], L. M. Hackel[4], D. R. Hagen[3], N. B. Jostmann[3], & D. M. Amodio*[1, 3]

[1]New York University
[2] VU University Amsterdam
[3]University of Amsterdam
[4]University of Southern California

*Corresponding author:

David M. Amodio
Department of Psychology
University of Amsterdam
david.amodio@gmail.com

Keywords: Stereotypes, Prejudice, Learning, Instrumental, Computational

Abstract

How are societal stereotypes transmitted to individual-level group preferences? We propose that exposure to a stereotype, regardless of whether one agrees with it, can shape how one experiences and learns from interactions with members of the stereotyped group, such that it induces individual-level prejudice—a process involving the interplay of semantic knowledge and instrumental learning. In a series of experiments, participants interacted with players from two groups, described with either positive or negative stereotypes, in a reinforcement learning task presented as a money sharing game. Although players' actual sharing rates were equated between groups, participants formed more positive reward associations with players from positively-stereotyped than negatively-stereotyped groups. This effect persisted even when stereotypes were described as unreliable and participants were instructed to ignore them. Computational modeling revealed that this preference was due to stereotype effects on priors regarding group members' behavior as well as the learning rates through which reward associations were updated in response to player feedback. We then show that these stereotype-induced preferences, once formed, spread unwittingly to others who observe these interactions, illustrating a pathway through which stereotypes may be transmitted and propagated between society and individuals. By identifying a mechanism through which stereotype knowledge can bypass explicit beliefs to induce prejudice, via the interplay of semantic and instrumental learning processes, these findings illuminate the impact of stereotype messages on the formation and propagation of individual-level prejudice.

Significance Statement

How do social stereotypes that exist in society transform into individual-level prejudices? In a series of experiments, we show that stereotype exposure shapes how we learn about group members in direct social interactions, and that this learning bias predicts the formation of group preferences. We further show that, once learned, these group preferences are transmitted to naïve observers who merely witness interactions between stereotyped group members and a person with stereotype knowledge. Finally, we show that this pattern of prejudice formation and propagation occurs even when people view the stereotype as unreliable and attempt to inhibit its influence. Together, these studies reveal a mechanism through which stereotypes may be transmitted and propagated between society and individuals.

How do explicit stereotypic messages about social groups become internalized in a individual's own preferences and behaviors? When a politician refers to a group as "criminals and rapists," as Donald Trump famously did during his 2015 campaign announcement, people may dismiss the epithets as mere rhetoric. Yet such messages may nevertheless be encoded in the listener's memory. We asked whether such knowledge, even when dismissed, can shape how people subsequently perceive and learn from members of the targeted group in direct interactions, such that it transforms into personal group preferences—a process representing the transmission of prejudice from societal-level stereotypes to individual-level attitudes.

To understand how stereotype knowledge may transform into individual-level prejudice through social interaction, we considered the interplay of learning mechanisms underlying stereotype knowledge and social-interactive impression formation (*1*–*3*). Stereotypes are societally-held beliefs about a group and its members, encoded in semantic memory (4–6). By providing expectancies for group members' behaviors, stereotypes can shape how we perceive and interpret a person's actions (7–10). However, like other forms of semantic knowledge, mere knowledge of a stereotype does not imply its endorsement: most low-prejudice individuals explicitly reject social stereotypes and inhibit stereotype effects on their judgments and behaviors (8, 11–13). This longstanding view within intergroup bias research suggests that an individual's personal beliefs are insulated from their knowledge of societal stereotypes (10, 14). From this perspective, exposure to a stereotype message should not, by itself, induce individual-level prejudice.

Here, however, we considered an unexplored possibility: If stereotypes provide expectancies for a group member's behavior, can stereotype knowledge inadvertently bias how we experience and learn about group members during direct social interactions? In direct interactions, a perceiver learns about a group member through the exchange of action and feedback—a process characterized by instrumental learning (i.e., reward reinforcement, 1, 3, 15). In contrast to stereotype knowledge, represented by semantic concepts, instrumental learning forms incrementally through repeated interaction and feedback, encoded in terms of reward value, and is expressed in choice behaviors that reflect an individual's personal, internalized preferences (16–18). Furthermore, whereas stereotype knowledge is explicit and easily inhibited in overt responses, instrumental learning is considered nondeclarative, such that it can form without explicit awareness of learning contingencies (19, 20). As a result, it may be especially difficult for a learner to detect or inhibit unwanted influences on the impressions they form of people through instrumental learning in direct interactions.

How might stereotypes influence instrumental learning? Instrumental learning can be shaped by priors, such as past experiences or knowledge, which can affect one's expectations about feedback and the degree to which a reward association is updated (17, 21). If stereotypes function as priors in instrumental learning, then exposure to a stereotype message may also bias reward expectancies associated with a group and the degree to which this reward association is updated in response to a group member's feedback, potentially inducing an internalized group-based preference. This process, involving the interplay of semantic and instrumental learning, would represent a pathway through which stereotype knowledge may bypass explicit egalitarian beliefs to produce individual-level prejudice.

Based on this analysis, we hypothesized that stereotype messages can induce personal group-based preferences through two concerted processes: First, exposure to a positive or negative stereotype sets initial expectations (i.e., *priors*) for a group member's behavior; second, stereotypes influence learning—that is, the degree to which reward representations are updated in response to feedback across repeated interactions (i.e., the *learning rate*)—such that updating occurs differently for members of positively and negatively stereotyped groups.

We tested this *stereotype learning* hypothesis across eight experiments in which we predicted that stereotype descriptions of groups would influence participants' instrumental learning during direct interactions with group members, even when participants explicitly dismiss the stereotype. We examined this effect in participants' behaviors and tested our hypothesis using computational modeling, and then further examined how such biases, once acquired and expressed, may spread to others who observe these direct interactions.

In experiments 1-3, participants interacted with people from two different social groups in an online point sharing game. These groups were labeled "Group A" and "Group B" (counterbalanced) in the task, ostensibly to maintain their anonymity, but described using positive or negative societal stereotypes associated with White and Black Americans, respectively (14). Group A was characterized as coming from a relatively wealthy, safe, and highly educated community, whereas Group B's community was characterized as relatively poor and uneducated and with a high crime rate (Figure 1a; see Supporting Information [SI]). This approach allowed us to isolate effects of stereotypes on learning while controlling for participants' existing group knowledge. Despite these group descriptions, participants were told that individual group members varied in their tendency to share points during the game and therefore, given participants' explicit goal to earn points, they should attend to the individual sharing rate of each player. Participants then completed a point sharing game with members of both groups, receiving cash payouts for their winnings.

The sharing game was adapted from a widely-used probabilistic reward reinforcement learning task (22). In this version, participants interacted with four players from each group. Within groups, each player shared points at a different fixed rate (70%, 60%, 40%, or 30%), but average sharing rates were equated between groups (Figure 1). Participants first completed a training phase, in which they could learn from feedback on each trial and, by choosing players who shared, earn points that would be converted to a cash bonus. On each round of training (160 trials), participants were presented with a preset pair of players—one from each group, with fixed complementary sharing rates (e.g., Players A and B)—and chose, via button press, with whom to interact (Figure 1). Reward feedback, displayed immediately beneath the image of the chosen player, indicated whether the chosen player shared (+1 or 0 points). Participants knew that only one player would share on each round.

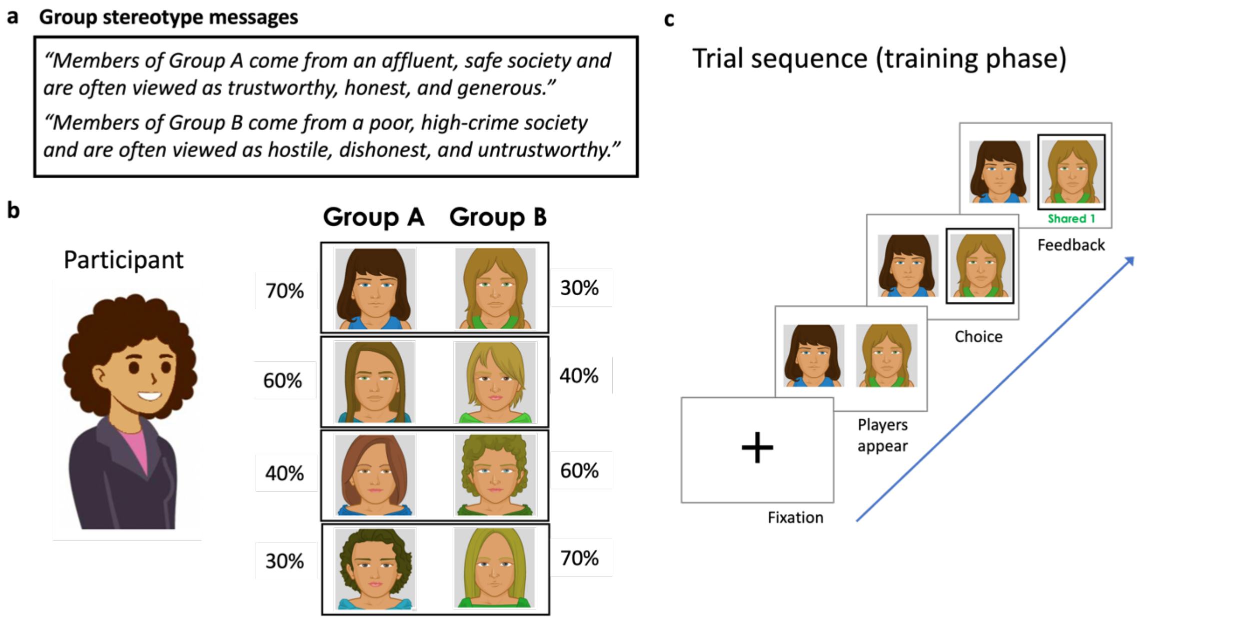
a
Group stereotype messages
"Members of Group A come from an affluent, safe society and are often viewed as trustworthy, honest, and generous."
"Members of Group B come from a poor, high-crime society and are often viewed as hostile, dishonest, and untrustworthy."
b
Participant
Group A
Group B
70%
30%
60%
40%
40%
60%
30%
70%
c
Trial sequence (training phase)
Shared 1
Feedback
Choice
Players appear
+
Fixation

Figure 1. Schematic of the learning task training phase. **a,** Participants were exposed to positive and negative stereotype messages regarding each group and then **b,** interacted with members of two groups. Group labels (A and B), member features (e.g., hair, shirt color), and gender were counterbalanced across participants. **c,** In the training phase, participants chose between players (group members) with reciprocal reward rate (e.g., 70% and 30%) and received reward feedback, as shown in this sample trial. In a test phase, participants chose between all possible intergroup pairs of players (e.g., 70% and 70%) and received no feedback.

Following the training phase, participants completed the test phase (96 trials), which provided a readout of their learning. In the test phase, participants viewed and selected between all possible pairs of Group A and B members. This allowed us to assess participants' choice preferences between novel pairs of players at every combination of reward rate. Hence, the test phase provided a fine-grained behavioral assessment of learned reward associations with each member of the two groups (22). Although feedback was not provided to prevent further learning, participants were told they would receive cash payout for their test phase choices following task completion.

Results

In **Study 1** (*N* = 61 laboratory participants), we tested whether stereotypic group descriptions influenced participants' choices of individual players, despite equivalent sharing rates between groups—the hallmark of group-based prejudice. Analysis of test phase behavior showed that while participants learned the general pattern of rewards, choosing players with higher sharing rates on average (*B* = 2.68, *SE* = 0.19, Wald *z* = 14.43, *p* < .001; all tests two-tailed), their choices were also significantly affected by players' group membership (*B* = 0.52, *SE* = 0.06, Wald *z* = 9.33, *p* < .001; Figure 2a). This effect of group membership emerged despite participants' extensive direct experience with players' actual sharing rates, which were equated between groups and thus contradicted the stereotypes, as well as the monetary incentive to choose accurately. These results revealed that choice preferences were guided by the group stereotype as well as actual reward feedback.

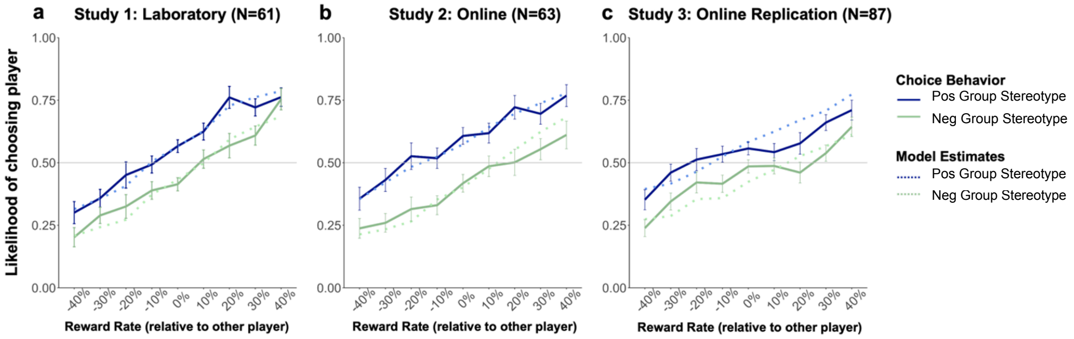

Figure 2. Behavioral choice preferences during the test phase in Studies 1–3 as a function of reward rate and group stereotype (Panels a–c, respectively). Participants' choices (solid lines) demonstrated both successful learning of rewards and a group bias. Reward rate (x axis) represents the actual reward rate of a given player minus the actual reward rate of the alternative player in a trial. Error bars indicate standard error. Dotted lines show estimates simlulated from the stereotype-learning model, which combined group-based priors and separate learning rates.

Next, to test our specific hypothesis that this effect involved the influence of stereotype knowledge on instrumental learning, we fit behavior to a computational model specifying this process, adapted from (23). We conceptualized stereotype effects on group expectancy as separate *priors* for positively and negatively stereotyped groups, which set participants' initial choice tendencies. Stereotype effects on learning (i.e., the updating of reward associations) were represented by separate *learning rates* for positively- and negatively-stereotyped groups. Thus, according to this hypothesized *stereotype learning model* (Figure 3), the behavioral effects of stereotypes on instrumental learning reflect a combination of divergent group priors and separate group learning rates.

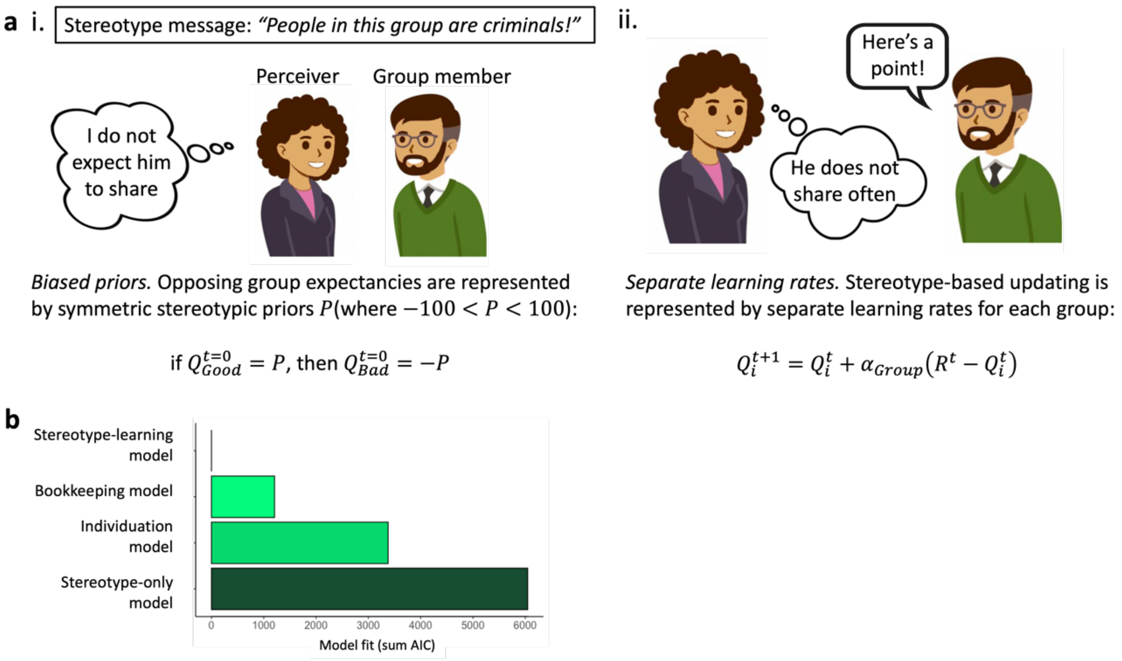
a
i.
Stereotype message: *"People in this group are criminals!"*
Perceiver
Group member
I do not expect him to share
Biased priors. Opposing group expectancies are represented by symmetric stereotypic priors P(where $-100 < P < 100$):
if $Q^{t=0}_{Good} = P$, then $Q^{t=0}_{Bad} = -P$
ii.
Here's a point!
He does not share often
Separate learning rates. Stereotype-based updating is represented by separate learning rates for each group:
$Q^{t+1}_i = Q^t_i + \alpha_{Group}(R^t - Q^t_i)$
b
Stereotype-learning model
Bookkeeping model
Individuation model
Stereotype-only model
0
1000
2000
3000
4000
5000
6000
Model fit (sum AIC)

Figure 3. a, According to the stereotype-learning model, (i) a stereotype message creates a positive or negative expectancy (prior) for a group member's behavior, and (ii) in subsequent interactions, perceivers update the value of positively- and negatively-stereotyped group members with separate learning rates. **b,** Model comparison (shown for Study 1) indicated the stereotype-learning model fit best to data compared with other plausible models of stereotyping and impression formation.

We compared the stereotype learning model with alternatives representing existing models of stereotyping and impression formation: (a) a *bookkeeping model* (e.g., 24, 25), in which new learning incrementally replaces the stereotype (biased priors and a single, unbiased learning rate), (b) an *individuation model*, in which learning is based only on players' actual behavior (a single learning rate and no priors), and (c) a *classic stereotyping model*, in which stereotypes determine responses without learning (biased priors with no learning), in addition to other plausible reinforcement learning and Bayesian accounts (see Method and SI for model specifications and results). Model comparisons indicated that the stereotype learning model, which included stereotype priors and separate group learning rates, was most consistent with observed behavior, supporting our hypothesis (Figure 3b; model fits in Table S2).

This effect was replicated in two online experiments (**Study 2**: N = 62; **Study 3**: N = 87): In both, stereotypic group descriptions again significantly influenced participants' test phase choice preferences (Study 2: B = 0.79, SE = 0.06, Wald z = 13.86, p < .001; Study 3: B = 0.48, SE = 0.05, Wald z = 9.58, p < .001), in addition to player's actual reward rates (Figure 1b and c; see SI). Again, this group bias emerged despite equivalent average reward rates between groups, participants' explicit goal to individuate, and the financial incentive to choose players based on their actual behavior.

Computational modeling of Study 2 and 3 data each replicated the results of Study 1, such that choice behavior was most consistent with a model that included group-based priors and separate group learning rates (see SI). Using combined data from Studies 1-3, parameter estimates of priors and group-specific learning rates, derived from the stereotype learning model, were submitted to a regression predicting group-based choice behaviors. Results indicated that the group bias in preferences reflected stereotype-based priors as well as insufficient updating for the negatively-stereotyped group; that is, initial expectancies for the negatively-stereotyped group were lower, relative to the positive group, and were not sufficiently updated in response to group members' actual reward feedback (see SI).

Study 3 was designed to address three additional aims. The first was to establish that stereotype descriptions were encoded in semantic memory. Participants completed a task in which they sorted stereotype traits used in the group descriptions to corresponding group labels. Classification accuracy for group stereotypes was significantly greater than chance (M = 75.02%; t = 7.87, 95% CI[0.68;0.78], df = 74, p < .001), indicating that stereotype descriptions were indeed encoded in memory.

The second aim was to test whether participants were aware of the stereotype effect on their choice preferences. To this end, we assessed participants' subjective estimates of player sharing rates following the task. The subjective estimates were significantly predicted by the group stereotype, B = 31.31, SE = 8.49, t = 3.69, p < .001, independently of players' actual sharing rates, suggesting that players misperceived a group difference in sharing (when none actually existed). However, when this subjective misperception was covaried in an analysis of choice behavior, the effect of group stereotype remained significant, B = 0.21, SE = 0.05, t = 3.97, p < .001. Thus, the effect of stereotypes on instrumental choice preferences was largely implicit.

The third aim was to determine whether participants could inhibit the influence of stereotypes in their explicit responses, despite the stereotype effect on instrumental learning. Following the main task, Study 3 participants completed a single-round trust game with each player, in which they could entrust a portion of their winnings from the sharing game to a player for a potentially larger return (26, see SI). Participants were told that the entrusted amount would be quadrupled, and that the return from each player would be based on that players' responses in the prior sharing game. Unlike decisions in the choice task, which involved binary classifications made under a 2 s response deadline, trust game decisions involved deliberation about potential payouts, with 10 choice options per round and unlimited decision time. Results showed that participants' explicit trust decisions reflected only the players' actual reward rates from the sharing game, with more money entrusted to higher-reward players, *B* = 5.87, *SE* = 1.34, *t*(693) = 4.38, *p* < .001, 95% CI [3.24, 8.50]. Trust decisions were not influenced by group stereotypes, *B* = 0.63, *SE* = 0.42, *t*(693) = 1.48, *p* = .14, 95% CI [-0.19, 1.45], suggesting that the stereotype knowledge was successfully inhibited in explicit responses.

Finally, to ensure that the group effects on choice preferences in Studies 1-3 were not due to wealth cues included in the stereotypes, this procedure was repeated in **Study 4** *(N* = 105, preregistered: https://aspredicted.org/RBP_FXD), using stereotype descriptions that omitted references to wealth. Study 4 results replicated those of Studies 1-3: Participants' behavioral choice preferences again reflected group stereotypes (*B* = 0.36, *SE* = 0.04, Wald *z* = 8.46, *p* < .001), in addition to players' actual reward rates (*B* = 2.29, *SE* = 0.14, Wald *z* = 16.76, *p* < .001), demonstrating that the stereotype effect on instrumental preferences was not due to beliefs about a player's wealth. Moreover, as in Study 3, participants self-reported a group difference in sharing that did not actually exist, *B* = 4.44, *SE* = 1.34, *t* = 3.32, *p* < .001—a misperception suggesting they believed that their group preference was driven by players' actual behavior (see SI).

Together, Studies 1-4 demonstrate that exposure to explicit social stereotypes leads to the formation of internalized group preferences through the process of instrumental learning during interactions with group members. Computational modeling indicated that this pattern reflects the influence of stereotypes on initial expectancies (priors) as well as updating of group member preferences based on reward feedback (leaning rates). This effect of stereotypes on instrumental learning appeared to be implicit; whereas participants inhibited stereotype effects in their explicit decisions, they were unaware of the stereotype's influence in the instrumental choice preferences.

Having observed the transmission of societal stereotypes to individual-level group preferences in Studies 1-4, we next considered a secondary form of transmission, whereby stereotype-based preferences spread to people who merely observe interactions between a stereotype-exposed actor and group member (27). Prior research shows that observers often misattribute an actor's biased behaviors to qualities of the group member, leading the observer to form their own group bias (27, 28). These findings suggest a pathway through which societal-level stereotypes, once internalized in an individual's group preferences, may propagate back into a society.

In **Study 5** (*N* = 124, preregistered: https://aspredicted.org/STK_EXP), participants played the money sharing game as in Studies 1-4. However, instead of learning directly from group members in a training phase, participants observed the training-phase choices and feedback of a prior participant (demonstrator) across 160 trials. Observers were told they should observe and learn from each player's feedback to improve their own chances of winning money in a subsequent test phase with the same players. Crucially, observers were not exposed to the stereotype descriptions provided to demonstrators; they were told only that players came from two different groups. Each Study 5 participant (observer) viewed the learning phase interactions of a participant from Study 2, in which a demonstrator made choices and received feedback from players. Two observers were yoked to each Study 2 direct learner. Participants then made their own choices in a test phase (identical to the test phase in Studies 1-4). Following the task, participants reported estimated reward rates for each player. This yoked design allowed us to trace the influence of the stereotype message through the direct learner to the group preferences of an observer.

Did the mere observation of demonstrators' behavior and feedback induce a group preference in observers? It did: observers exhibited a significant group bias in their own test phase choices, despite having no exposure to the stereotype (*B* = 0.32, *SE* = 0.04, Wald *z* = 8.03, *p* < .001), in addition to learning from players' rewards (*B* = 1.49, *SE* = 0.09, Wald *z* = 16.73, *p* < .001, Figure 5). Moreover, the magnitude of their group bias correlated with the degree of bias exhibited in the demonstrator's own test phase choices (*B* = 0.28, SE = 0.09, Wald *z* = 3.21, *p* = .001), indicating that the demonstrator's degree of prejudice was transmitted to the observer. These findings suggest a cycle of bias propagation, from societal stereotypes to an individual's group preferences, and then to naïve third-party observers.

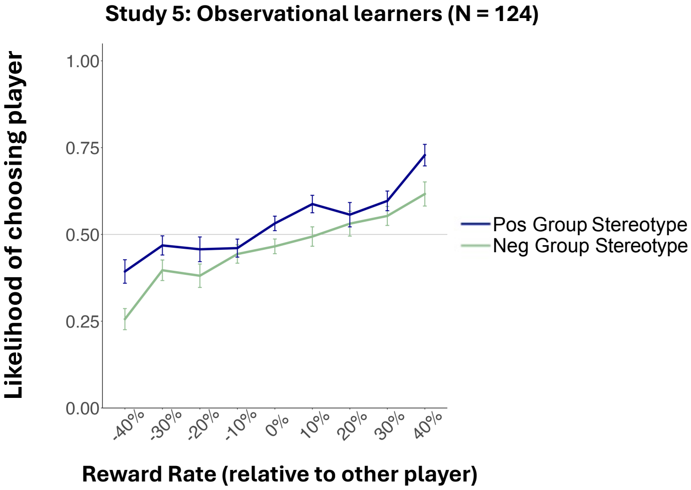

Figure 4. Behavioral choice preferences during the test phase for observational learners in Study 5 as a function of reward rate and group stereotype. Choice preferences of naïve observers reflected the stereotype-biased preferences of demonstrators, in addition to players' actual reward rates. The x-axis represents the difference between actual reward rates of the two players on a given trial.

Finally, having found that social stereotypes can be internalized in one's own choice preferences through instrumental learning and propagated to others through observation, we returned to the question we began with: can exposure to societal stereotypes induce internalized group preferences through social-instrumental learning even when people explicitly attempt to ignore the stereotype?

In **Study 6** ($N = 106$, https://aspredicted.org/BDH_CDH), participants were exposed to group stereotypes as in Study 4. However, unlike prior studies, these participants were then informed that (a) the descriptions were common stereotypes which were unreliable and (b) participants should attend only to the feedback of individual players to maximize points. This procedure mimicked the common real-world experience of being exposed to stereotype information but cautioned to ignore it. Nevertheless, despite these instructions, participants' choice behavior continued to reflect the group stereotypes, $B = 0.57$, $SE = 0.04$, Wald $z = 13.65$, $p < .001$, in addition to players' actual reward rates, $B = 2.33$, $SE = 0.13$, Wald $z = 17.50$, $p < .001$. Moreover, participants' self-reports of player sharing rates were predicted by group membership, $B = 4.83$, $SE = 1.41$, $t = 3.42$, $p < .001$, in addition to their actual reward rates, $B = 46.05$, $SE = 4.46$, $t = 10.33$, $p < .001$, again suggesting that the stereotypes led participants to misperceive a difference in group members' behavior that did not actually exist.

Study 7 ($N = 154$, https://aspredicted.org/V8W_7ZC) repeated the Study 6 procedure with more stringent instructions: After viewing group stereotypes and receiving instructions to individuate, but before beginning the main task, participants completed an understanding quiz. This quiz required participants to correctly indicate their task goal—to choose based on individual player feedback and not group stereotypes—before proceeding to the main task. Despite these explicit instructions and confirmation of participants' understanding, participants' choice preferences continued to reflect the stereotype messages ($B = 0.44$, $SE = 0.04$, Wald $z = 12.50$, $p < .001$), in addition to players' actual rewards ($B = 2.33$, $SE = 0.11$, Wald $z = 20.90$, $p < .001$, Figure 5a). Furthermore, participants' self-reported estimates of player sharing rates were predicted by group membership ($B = 1.98$, $SE = 0.18$, $t = 11.20$, $p < .001$), in addition to actual reward rates ($B = 41.02$, $SE = 0.55$, $t = 75.1$, $p < .001$). Thus, as in Study 6, participants were unable prevent the influence of stereotypes on their instrumental learning of group members, and they again misperceived a group difference in player sharing rates that did not actually exist.

In a final study, we tested whether the hypothesized cycle of bias transmission—from societal stereotype to individual to community members—would emerge even when direct learners dismissed the stereotype. In **Study 8** ($N = 154$, https://aspredicted.org/H6M_SSZ) participants observed the learning phase trials of Study 7 participants—direct learners who were instructed to ignore the stereotype. Observers, naïve to the stereotype messages, were matched to Study

7 demonstrators in a yoked design (1-to-1 yoking), similar to Study 5. Here again, we found that observers formed group preferences that were consistent with stereotype knowledge of demonstrators (*B* = 0.19, *SE* = 0.03, Wald *z* = 5.67, *p* < .001), in addition to players' actual reward feedback (*B* = 1.50, *SE* = 0.11, Wald *z* = 13.96, *p* < .001, Figure 5b). The degree of group preference acquired by observers was directly associated with the preference of their respective demonstrator (*B* = 0.15, *SE* = 0.05, *t* = 2.30, *p* = .003). These results demonstrate that stereotype messages can induce a prejudice in direct learners which can then spread to naïve observers, even when the direct learners explicitly attempted to ignore the stereotype.

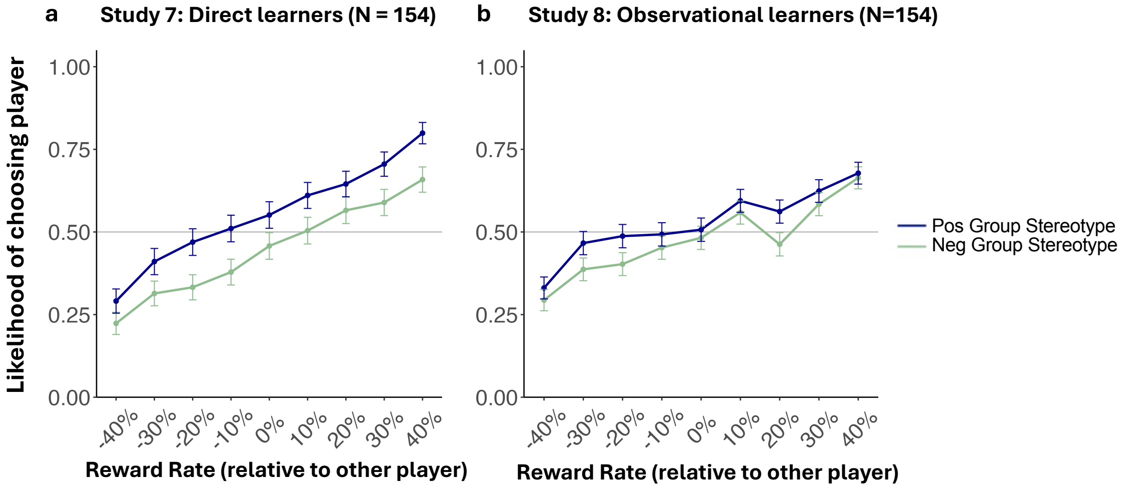

Figure 5. a, Behavioral choice preference for the test phase of Study 7. Participants' choices reflected the group stereotype, in addition to player reward rates, despite instruction to ignore stereotypes. **b,** In Study 8, observers naïve to group stereotypes who viewed the learning phase choices and reward feedback of Study 7 participants showed a group bias in their own test phase choice preferences. The x-axis represents the difference between actual reward rates of the two players on a given trial. Error bars indicate standard error.

Discussion

We asked whether exposure to societal stereotypes can induce personal group-based preferences by shaping the way one learns about group members in direct interactions. Across six studies, we found that positive and negative group stereotypes, conveyed explicitly, shaped the process of instrumental learning in direct interactions with group members. Computational modeling suggested this effect involved the interplay of two processes: stereotypes set initial expectancies for each group and then influenced the updating of reward values associated with individual group members. This effect of stereotype exposure on instrumental learning appeared to occur implicitly: although participants were aware of the stereotype content and could inhibit its effect in their explicit trust decisions, they could not prevent its effect on their instrumentally-learned preferences toward group members. These findings reveal a mechanism through which mere exposure to stereotype information can bypass an individual’s explicit intentions to induce an internalized group preference.

Next, to examine the broader impact of this mechanism for societal-level prejudice, we asked whether these group choice preferences—formed in response to stereotype exposure—could spread to observers of these interactions via social learning (27, 29). Indeed, in two additional studies, we found that stereotype-induced preferences in participants’ choice behavior were acquired unwittingly by observers who, after viewing this behavior with no knowledge of group stereotypes, expressed stereotype-consistent preferences in their own choices. These findings build on our initial results to illustrate how group preferences produced by stereotype exposure may propagate throughout a community.

This research introduces a model of intergroup bias that describes how exposure to a societal stereotype can induce individual-level prejudice, even among individuals who personally reject the stereotype. Although the importance of considering both individual and societal aspects of intergroup bias is well recognized (30–33), few studies have examined the psychological pathways through which they interact (34). By integrating existing models of stereotyping, based on semantic knowledge representations, with instrumental learning models of direct and observational learning, the present research specifies such a pathway. In doing so, it provides a theoretical framework for understanding how systemic disparities in one’s environment may be internalized in the mind of the individual.

The transmission of societal stereotypes to individual prejudice observed in our studies appeared to occur without participants’ awareness. That is, while participants were aware of the stereotype content and could inhibit its effect on their explicit responses, they appeared unaware of the stereotype influence on their interaction-based preferences formed through instrumental learning. This effect was likely due, in part, to its indirect nature: although participants’ explicit goal was to choose players based on individual sharing rates, the task afforded an indirect influence of group membership—much like in real intergroup interactions—which may have been difficult to detect and inhibit. This pattern is further consistent with the nondeclarative operation of instrumental learning which, in past research, has been shown to occur in the absence of awareness (19, 20). These features—the indirect nature of stereotypes on social-interactive instrumental learning and its nondeclarative operation—suggest a potent form of implicit prejudice that has not been previously explored.

A potential alternative account of our findings is that participants simply applied the stereotype knowledge they were given, much like a base rate. However, several aspects of our findings suggest that a “base rate” explanation is unlikely. First, computational modeling across six studies consistently showed that group preferences were explained not just by stereotype priors, but also by stereotype effects on learning; by contrast, a “base rate” model in which preferences were determined by stereotype priors without learning (the “stereotype only” model) was the worst-fitting model. Second, participants formed group preferences even when the stereotype was explicitly discounted and they were instructed to ignore it, and despite financial incentives opposing the stereotype. And third, participants reported perceiving a group difference in sharing despite equated reward rates, further suggesting that participants’ group preferences reflected their direct learning experiences and not merely the application of a base rate.

Our research contributes methodological advances to the study of intergroup bias through its use of computational modeling to systematically test and compare theories of stereotype function. Here, we adapted models of rule-based priors on reinforcement learning (21, 35) to address the effect of stereotype knowledge on interactive learning (36). By formalizing and comparing alternative models, we found strong support for the hypothesized *stereotype learning model*, whereby stereotypes operated as priors and differentially affected learning from group members. This approach complements prior research on biased sampling in the formation of prejudice (37–39), further illustrating how computational modeling may be used fruitfully to investigate mechanisms of social cognition and their interplay with features of society (27, 40–44).

More broadly, our findings show that messages promoting societal stereotypes are more than mere words; exposure to biased group descriptions can shape one’s subsequent experiences with members of the group, perhaps without one’s knowledge, in ways that confirm the message and spread it to others. This process—whereby societal stereotypes are transmitted to personal group preferences—may also help to explain how systemic biases, such as institutional inequality, may be transmitted via stereotypes from social structures to the minds of individuals (45–48). As society continues to grasp the impact of polarizing sociopolitical rhetoric, from campaign ads to social media, our findings suggest that its influence may be more potent and far-reaching than previously thought. Yet, by illuminating the processes through which explicit societal messages may induce personal bias in the individual, these results may inform new approaches to reducing their impact.

Materials and Methods

Stereotype manipulation. Upon starting the experiment, participants learned that they would play a money sharing game with players from two social groups. Before beginning the task, participants were given the following descriptions of these groups (counterbalanced across participants):

> *“In the main task you will play an interactive money-sharing game with people from two different groups who come from different places. For the purpose of this study, we will refer to these groups as Group A and Group B, and their members will be represented by*

> *avatars. Members of Group A live in a more affluent society, where crime is low and most people have good jobs. People from Group A are often perceived to be trustworthy, honest, and generous to others, and they are proud of their success. Group B, by comparison, lives in a society that is economically poor, with a high rate of unemployment and serious crimes such as robbery, assault, and murder. People from Group B are often perceived to be hostile, untrustworthy, and dishonest."*

Participants were then shown avatars representing players from each group, with color cues (blue vs. green clothing, darker vs. lighter hair) signaling group membership. Participants interacted with either all female or all male-appearing avatars. Participants were instructed that players had participated in a previous experiment in which they decided how many points (redeemable for a monetary bonus) to share. Participants were further told that different players shared different amounts, and they should learn who shared more often to win the most points.

Learning task. The main learning task consisted of a 160-trial training phase and a 96-trial test phase. In the training phase, participants always chose between two targets—one from each group—with reward probabilities adding up to 1 (70% vs. 30% or 60% vs. 40%). Although the reward feedback varied within groups, there was no difference between groups. On each trial, a face pair was shown for a maximum of 2 s, during which time a response was required. Reward feedback (+1 or 0 points) appeared immediately following choice. Player gender and group color cue (blue or green) were counterbalanced across participants, and player identity was randomized such that individual players were assigned to random reward rates for a given participant.

The test phase provided a readout of learned reward values. Participants chose between all combinations of targets from different groups, always with one Group A member and one Group B member. Each pair was shown for a maximum of 2 s, during which time a response was required, followed by a 1000 ms intertrial interval. Feedback was not given, to prevent further learning, but choices were nonetheless incentivized.

Computational modeling. Computational reinforcement learning (RL) models used to evaluate our hypothesis and alternatives were based on the standard Q-learning rule:

$$Q_i{}^{t+1} = Q_i{}^{t} + \alpha(R^t - Q_i{}^{t})$$

where Q_i is the action value of selecting option *i* in trial *t*, *R* is the reinforcement [no reward = 0, reward = 1] received in trial *t*, and α ($0 \le \alpha \le 1$) is a learning rate parameter, which determines how much the difference between the received and the predicted reinforcement (the prediction error) affects subsequent value estimates

These Q-values were then transformed into decision probabilities using a standard Softmax function:

$$P_i = \frac{e^{Q_i/\beta}}{\sum_{j=1}^{2} e^{Q_j/\beta}}$$

To examine effects of group-based initial expectations, the model was formulated using a symmetrical prior parameter (ranging from -100 to +100):

$$Q_{Good}^{t=0} = prior, Q_{Bad}^{t=0} = -prior$$

To examine effects of target group on learning, models included separate learning rates as a function of group membership:

$$Q_{i,group}{}^{t+1} = Q_{i,group}{}^{t} + \alpha_{group}(R^t - Q_{i,group}{}^{t})$$

Detailed descriptions of methods may be found in the Supporting Information.

Acknowledgments

Portions of this paper are reported in D. Schultner's PhD thesis. B.R.L. acknowledges support from a Wallenberg Academy Fellow grant from the Knut and Alice Wallenberg Foundation (KAW 2021.0148) and a Starting Grant (SOLAR ERC-2021-STG – 101042529) from the European Research Council.

Funding: This work was supported by the Netherlands Organisation for Scientific Research (VICI 016.185.058) to D.M.A.

References

1. D. M. Amodio, Social Cognition 2.0: An Interactive Memory Systems Account. *Trends Cogn. Sci.* **23**, 21–33 (2019).
2. T. E. J. Behrens, L. T. Hunt, M. F. S. Rushworth, The computation of social behavior. *Science* **324**, 1160–1164 (2009).
3. L. M. Hackel, B. B. Doll, D. M. Amodio, Instrumental learning of traits versus rewards: dissociable neural correlates and effects on choice. *Nat. Neurosci.* **18**, 1233–1235 (2015).
4. J. W. Sherman, Development and mental representation of stereotypes. *J. Pers. Soc. Psychol.* **70**, 1126–1141 (1996).
5. J. L. Hilton, W. von Hippel, STEREOTYPES. *Annu. Rev. Psychol.* **47**, 237–271 (1996).
6. S. T. Fiske, “Stereotyping, prejudice, and discrimination” in *The Handbook of Social Psychology, Vols*, D. T. Gilbert, Ed. (McGraw-Hill, Boston, 1998), pp. 1–2.
7. J. M. Darley, P. H. Gross, A hypothesis-confirming bias in labeling effects. *J. Pers. Soc. Psychol.* **44**, 20–33 (1983).
8. P. G. Devine, Stereotypes and prejudice: Their automatic and controlled components. *J. Pers. Soc. Psychol.* **56**, 5 (1989).
9. K. Kawakami, D. M. Amodio, K. Hugenberg, “Intergroup Perception and Cognition” in *Advances in Experimental Social Psychology*, Advances in experimental social psychology., (Elsevier, 2017), pp. 1–80.
10. Z. Kunda, S. J. Spencer, When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychol. Bull.* **129**, 522–544 (2003).
11. D. M. Amodio, The social neuroscience of intergroup relations. *Eur. Rev. Soc. Psychol.* **19**, 1–54 (2008).
12. M. J. Monteith, L. Ashburn-Nardo, C. I. Voils, A. M. Czopp, Putting the brakes on prejudice: On the development and operation of cues for control. *J. Pers. Soc. Psychol.* **83**, 1029–1050 (2002).
13. B. K. Payne, Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *J. Pers. Soc. Psychol.* **81**, 181–192 (2001).
14. P. G. Devine, A. J. Elliot, Are racial stereotypes really fading? The Princeton trilogy revisited. *Pers. Soc. Psychol. Bull.* **21**, 1139–1150 (1995).
15. B. Lindström, I. Selbing, T. Molapour, A. Olsson, Racial bias shapes social reinforcement learning. *Psychol. Sci.* **25**, 711–719 (2014).
16. N. D. Daw, J. P. O’Doherty, P. Dayan, B. Seymour, R. J. Dolan, Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).
17. M. R. Delgado, R. H. Frank, E. A. Phelps, Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* **8**, 1611–1618 (2005).
18. D. Shohamy, A. D. Wagner, Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron* **60**, 378–389 (2008).
19. B. J. Knowlton, J. A. Mangels, L. R. Squire, A neostriatal habit learning system in humans. *Science* **273**, 1399–1402 (1996).
20. P. J. Reber, L. R. Squire, Parallel brain systems for learning with and without awareness. *Learn. Mem.* **1**, 217–229 (1994).
21. B. B. Doll, W. J. Jacobs, A. G. Sanfey, M. J. Frank, Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. *Brain Res.* **1299**, 74–94 (2009).
22. M. J. Frank, L. C. Seeberger, R. C. O’reilly, By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* **306**, 1940–1943 (2004).
23. M. J. Frank, B. B. Doll, J. Oas-Terpstra, F. Moreno, Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat. Neurosci.* **12**, 1062–1068 (2009).

24. M. Rothbart, "Memory processes and social beliefs" in *Cognitive Processes in Stereotyping and Intergroup Behavior*, D. L. Hamilton, Ed. (1981), pp. 145–181.

25. R. Weber, J. Crocker, Cognitive processes in the revision of stereotypic beliefs. *J. Pers. Soc. Psychol.* **45**, 961–977 (1983).

26. I. Bohnet, R. Zeckhauser, Trust, risk and betrayal. *J. Econ. Behav. Organ.* **55**, 467–484 (2004).

27. D. T. Schultner, B. R. Lindström, M. Cikara, D. M. Amodio, Transmission of social bias through observational learning. *Sci. Adv.* **10**, eadk2030 (2024).

28. M. Weisbuch, K. Pauker, N. Ambady, The subtle transmission of race bias via televised nonverbal behavior. *Science* **326**, 1711–1714 (2009).

29. A. Olsson, E. Knapska, B. Lindström, The neural and computational systems of social learning. *Nat. Rev. Neurosci.* **21**, 197–212 (2020).

30. G. W. Allport, *The Nature of Prejudice* (Addison-Wesley Publishing Company, 1954).

31. M. R. Banaji, S. T. Fiske, D. S. Massey, Systemic racism: individuals and interactions, institutions and society. *Cogn. Res. Princ. Implic.* **6**, 82 (2021).

32. F. L. Jones, Ethnic diversity and national identity. *Aust. N. Z. J. Sociol.* **33**, 285–305 (1997).

33. J. Sidanius, F. Pratto, "Social dominance theory: A new synthesis" in *Social Dominance*, (Cambridge University Press, 1999), pp. 31–58.

34. A. L. Skinner-Dorkenoo, M. George, J. E. Wages 3rd, S. Sánchez, S. P. Perry, A systemic approach to the psychology of racial bias within individuals and society. *Nat Rev Psychol* 1–15 (2023).

35. D. S. Fareri, L. J. Chang, M. R. Delgado, Computational substrates of social value in interpersonal collaboration. *J. Neurosci.* **35**, 8170–8180 (2015).

36. D. M. Amodio, M. Cikara, The social neuroscience of prejudice. *Annu. Rev. Psychol.* **72**, 439–469 (2021).

37. S. Allidina, W. A. Cunningham, Avoidance begets avoidance: A computational account of negative stereotype persistence. *J. Exp. Psychol. Gen.* **150**, 2078–2099 (2021).

38. X. Bai, S. T. Fiske, T. L. Griffiths, Globally Inaccurate Stereotypes Can Result From Locally Adaptive Exploration. *Psychol. Sci.* **33**, 671–684 (2022).

39. R. H. Fazio, J. R. Eiser, N. J. Shook, Attitude formation through exploration: valence asymmetries. *J. Pers. Soc. Psychol.* **87**, 293–311 (2004).

40. O. FeldmanHall, M. R. Nassar, The computational challenge of social learning. *Trends Cogn. Sci.* **25**, 1045–1057 (2021).

41. K. Kobayashi, J. W. Kable, M. Hsu, A. C. Jenkins, Neural representations of others' traits predict social decisions. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2116944119 (2022).

42. Y. Zhou, *et al.*, Learning from Ingroup Experiences Changes Intergroup Impressions. *J. Neurosci.* **42**, 6931–6945 (2022).

43. L. M. Hackel, D. Kogon, D. M. Amodio, W. Wood, Group value learned through interactions with members: A reinforcement learning account. *J. Exp. Soc. Psychol.* **99**, 104267 (2022).

44. I. J. Traast, D. Schultner, B. Doosje, D. Amodio, Race effects on impression formation in social interaction: An instrumental learning account. (2023).

45. M. M. Berkebile-Weinberg, A. R. Krosch, D. M. Amodio, Economic scarcity increases racial stereotyping in beliefs and face representation. *J. Exp. Soc. Psychol.* **102**, 104354 (2022).

46. A. R. Krosch, D. M. Amodio, Economic scarcity alters the perception of race. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 9079–9084 (2014).

47. B. K. Payne, H. A. Vuletich, K. B. Lundberg, The Bias of Crowds: How Implicit Bias Bridges Personal and Systemic Prejudice. *Psychol. Inq.* **28**, 233–248 (2017).

48. M. Vlasceanu, D. M. Amodio, Propagation of societal gender inequality by internet search algorithms. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2204529119 (2022).

Supporting Information

Transmission of societal stereotypes to individual-level prejudices through social instrumental learning

Schultner, D. T., Stillerman, B. S., Lindström, B. R., Hackel, L. M.,
Hagen, D. R., Jostmann, N. B., & Amodio, D. M.

Table of Contents

Study 1

Overview. In Study 1, we tested whether explicit descriptions of groups would bias participants' choices of who to interact with and win money from. Participants read descriptions of two fictional groups and then played an economic game with ostensible players from those groups. Group membership was, on average, not associated with reward probability and thus not beneficial cue for choice performance.

Method

Participants

Sixty-nine students at University of Amsterdam received course credit for their participation as well as a performance-based monetary bonus, ranging from $1.30 – $1.70. In this and subsequent reported studies, we excluded participants who failed to reach a learning criterion of 50% accuracy for 30%-70% player pairs during the test phase (i.e., A-B and G-H; see below for details of test phase procedure). In Study 1, this exclusion criterion yielded a final sample size of N = 61 (45 women, 16 men; M_{age} = 21.56 years, SD_{age} = 5.20 years).

Ethics approval was obtained from the human subjects institutional review board at the University of Amsterdam.

Procedure

Introduction and manipulation. Upon arrival to the lab and following informed consent, participants learned that they would play a money sharing game with players from two social-geographical groups. Before beginning the task, participants were given the following descriptions of these groups (with descriptions of Group A and B counterbalanced across participants):

> "In the main task you will play an interactive money-sharing game with people from two different groups who come from different places. For the purpose of this study, we will refer to these groups as Group A and Group B, and their members will be represented by avatars. Members of Group A live in a more affluent society, where crime is low and most people have good jobs. People from Group A are often perceived to be trustworthy, honest, and generous to others, and they are proud of their success. Group B, by comparison, lives in a society that is economically poor, with a high rate of unemployment and serious crimes such as robbery, assault, and murder. People from Group B are often perceived to be hostile, untrustworthy, and dishonest."

These descriptions were based on common societal stereotypes of White and Black Americans, respectively (1), and which also correspond to common stereotypes to White (native) and Moroccan Dutch immigrants.

This stereotype information was followed by a note that, despite these generalizations, there is individual variability, and that the participant should pay attention to individual players' behavior:

> "So, as you see, these groups are different in many ways. However, individuals within each group vary, too. You will need to learn about these people as you engage in repeated interactions in the task."

Participants were then shown avatars representing players from each group, with color cues (blue vs. green clothing, darker vs. lighter hair) signaling group membership (all other features were matched between groups). Participants were assigned, in counterbalanced fashion, to view either all female or all male-appearing avatars (Figure S1), to control for potential target gender effects. Participants were instructed that these players had participated in a previous experiment in which they decided how many points (redeemable for a monetary bonus) to share. Participants were further told that different players shared different amounts, and they should learn who shared more often to win the most points.

Categorization task. To ensure that participants learned the group identity of each player, they completed a categorization task embedded in a standard 7-block implicit association test (IAT). The first and fifth blocks of this IAT required simple categorizations of player to their group category, with accuracy feedback. This IAT was repeated at the very end of the task. Although not the focus of these studies, Time 1 IAT data indicated that explicit group descriptions alone created a significant IAT effect, with preference expressed toward Group A, similar to much prior research (e.g., 2), showing that IAT scores can be driven by a variety of influences including novel explicit group beliefs. The same pattern of Group A preference was observed at Time 2. We did not analyze IAT scores further, and the measure was dropped from all subsequent studies.

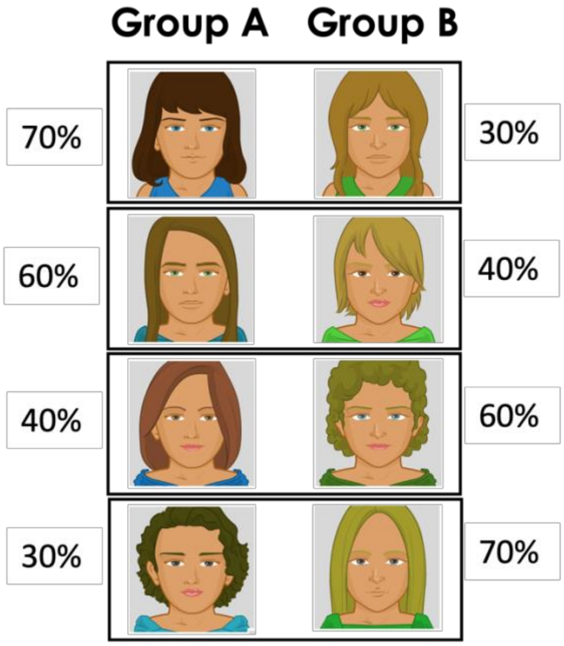

Figure S1. Sample avatars and schematic of reward probabilities.

Learning task. Next, participants completed the main learning task, which included a training phase of 160 trials and a test phase of 96 trials. In the training phase, participants always chose between two targets—one from each group—with complementary fixed reward probabilities (e.g., player pairs A-B, C-D, E-F, and G-H, see Figure S1). On each trial, a face pair was shown for a maximum of 2 s, during which time a response was required. Reward feedback (+1 or 0 points) appeared immediately following choice, and points were converted to a cash bonus at the task conclusion. The cover story was that this feedback was derived from past participants' actual choices of how often to share points. Crucially, although the reward feedback varied by individuals within groups, average reward rate for both groups was equated. Player gender and group color cue (blue or green) were counterbalanced across participants, and player identity was randomized such that individual players were assigned to random reward rates for a given participant. To win as much money as possible, participants were motivated to learn which target players tended to reward more often than others.

Next, in the test phase, in order to obtain a readout of learned reward values, participants chose between all combinations of targets from different groups (e.g., A-B, A-D, A-F, A-H, C-B, etc.), always with one Group A member and one Group B member. Each pair was shown for a maximum of 2 s, during which time a response was required, followed by a 1000 ms intertrial interval. Feedback was not given, to prevent further learning, but participants were told that correct choices would still be rewarded and paid out in the bonus at the end of the task.

Results

Primary analysis focused on test phase data and involved two approaches: multilevel regression and computational modeling. We detail the regression approach below and, for all studies, report computational modeling results in the section "Computational modeling".

Multilevel regression was used to test effects of (a) players' actual reward rate and (b) group membership on choice. Trials in which choices were made faster that 200 ms or slower than 2000 ms were excluded from analysis. Participants' trial-level choice data were submitted to a general linear mixed model predicting the likelihood that participants chose a given target player, nested by participant, with a logit link function. The primary model included by-participant random intercepts and the following predictors as fixed effects: players' actual reward rate, players' group membership, and their interaction. For completeness, we also report models with by-participant random slopes for the fixed effects. Sharing rates were equated between groups (t-test for a group difference: t = -0.45, df = 9505, p = 0.66).

Results indicated a significant effect of player's relative reward rate on choice, demonstrating learning of player reward rates, B = 2.68, SE = 0.19, Wald z = 14.43, p < .001. An examination of raw choice behavior revealed a relatively accurate mapping between participants' choices and the actual reward contingencies. Importantly, the effect of group membership on choice was also significant, such that participants were more likely to choose Group A members over Group B members, B = 0.52, SE = 0.06, Wald z = 9.33, p < .001. Indeed, when faced with two equally rewarding players, participants chose the Group A member 25% more often. The Reward Rate x Group interaction was not significant, B = -0.003, SE = 0.26, Wald z = -0.01, p = .992. The pattern was qualitatively identical in the random slopes model (Reward rates: B = 3.27, SE = 0.36, Wald z = 9.08, p < .001, Group bias: B = 0.65, SE = 0.23, Wald z = 2.79, p = .005)

To corroborate these findings, training phase data were submitted to the same general linear mixed model. Results replicated those of the test phase data, with evidence of accurate learning of player reward rates, *B* = 1.77, *SE* = 0.10, Wald *z* = 18.43, *p* < .001, as well as a bias to choose Group A, *B* = 0.29, *SE* = 0.04, Wald *z* = 6.95, *p* < .001. The reward rate by group interaction was marginally significant, *B* = -0.25, *SE* = 0.13, Wald *z* = -1.83, *p* = .068.

Study 2

Overview. In Study 2, we sought to replicate the findings of Study 1 in a new sample.

The procedure was identical to that of Study 1, except that the initial categorization task and IATs were dropped and it was conducted online rather than in the lab.

Ethics approval was obtained from the human subjects institutional review board at New York University.

Method

Participants. Participants were 78 Amazon Mechanical Turk (MTurk) workers (demographics unavailable due to technical error) who received $2.00 for their participation as well as a performance-based monetary bonus, ranging from approximately $0.30 – $0.40, derived from points earned during the task with a conversion rate of 2 points per cent. Participants who failed to reach a learning criterion of 50% accuracy for choices between 30% vs. 70% reward player pairs during the test phase (i.e., A-B and G-H; *N* = 16) were excluded. After exclusions, Study 2 had a final sample size of *N* = 62.

Procedure. Participants completed an online learning task, nearly identical to the one described in Study 1. Besides a slightly different look and feel for the online version, the only difference was that the groups described with positive and negative stereotypes were counterbalanced (e.g., whether Group A or Group B was described as the good group). For ease of reporting and visualizing the analysis, we refer to the group described with positive and negative stereotypes as “Group A” and “Group B,” respectively, in the results.

Results

Our analytical approach followed that of Study 1, with a focus on test phase choice data. Results indicated significant effects for players’ actual reward rate, demonstrating strong learning, *B* = 2.55, SE = 0.19, Wald *z* = 13.51, *p* < .001, and for group membership, such that participants strongly preferred Group A members independent of actual reward rates, *B* = 0.79, SE = 0.06, Wald *z* = 13.86, *p* < .001. The Reward Rate x Group Membership interaction was not significant, *B* = -0.35, SE = 0.26, Wald *z* = -1.32, *p* = .187. As in Study 1, the qualitative pattern was identical in the random slopes model Reward rates: *B* = 3.03, SE = 0.39, Wald *z* = 7.7, *p* < .001, Group bias: *B* = 1.02, SE = 0.37, Wald *z* = 2.8, *p* = .005)

As in Study 1, to corroborate these findings, we submitted the training phase data to the same general linear mixed model. The results replicated those of the test phase data, with significant effects of actual reward rate, *B* = 1.40, *SE* = 0.10, Wald *z* = 14.66, *p* < .001, and of group membership, evidencing a preference for Group A members, *B* = 0.45, *SE* = 0.04, Wald *z* = 10.59, *p* < .001. The interaction was not significant, *B* = 0.17, *SE* = 0.14, Wald *z* = 1.24, *p* = .214.

Study 3

Overview. In Study 3, we extended the procedure used in Studies 1 and 2 to include two additional post-learning-task measures: explicit beliefs of player reward rates and a trust game.

Method

Participants. Participants were 158 Amazon Mechanical Turk (MTurk) workers who received $2.00 for their participation as well as a performance-based monetary bonus, ranging from approximately $0.30 – $0.40, derived from points earned during the task with a conversion rate of 2 points per cent. We excluded participants who responded without variation for either of the post-learning-task measures (*N* = 18) and participants who failed to reach a learning criterion of 50% accuracy for 30%-70% player pairs during the test phase (*N* = 47). One participant was excluded due to a technical error resulting in invalid post-learning-task measures. Our exclusion of trials with invalid reaction times resulted in 5 participants being excluded altogether. These exclusions resulted in a final sample size of *N* = 87 (44 men, 36 women, 7 unreported; M_{age} = 34.9 years, SD_{age} = 10.0 years).

Ethics approval was obtained from the human subjects institutional review board at New York University.

Procedure. After reading the group instructions, participants completed a categorization task to reinforce the group membership of target players. Unlike the categorization task used in Study 1, which was embedded within an IAT, Study 3 used a stand-alone task that included the classification of both group member faces and trait terms that had been conveyed in the group description manipulation. Hence, participants were presented with pictures of players and stereotype words associated with the group descriptions (e.g., “wealthy,” “uneducated,” “trustworthy”) and classified each according to group label. They then completed the learning task, as in Study 2; however, in Study 3, Group A always associated with the positive stereotype description and Group B was associated with the negative stereotype description. This decision was made because A and B are often associated with better and worse options, respectively, and this could contribute to noise or confusion with the manipulation.

After the learning task, participants completed a subjective reward measure, in which they were asked, for each player in randomized order, “How many times out of a hundred would this player share with you?” For each player, participants typed their estimate of the player’s sharing rate, from 1 to 100, in a text box. This form of response was designed to assess declarative semantic knowledge, which might be expressed independent of any striatally-based instrumental tendencies that could influence responses on a slider-type scale (Knowlton et al. 1996).

Finally, participants played a single-shot trust game with each target player. They were told they had a 20-point pool and they could choose how much to invest in each player as a trustee. The trustee’s point amount would then be tripled and they could share any amount back to the participant. For each player, the participant selected a number of points to share, from 0 to 20, with options at 2-point intervals. Unlike the sharing game, which permitted a maximum of 2 seconds for binary decisions, the trust game allowed for deliberate choices with unlimited time and ten answer options per round.

Results

Choice behavior. Our analytical approach followed that of Studies 1 and 2. Multilevel regression predicting player choice again produced significant effects of players’ actual reward rate, *B* = 1.76, SE = 0.16, Wald *z* = 11.15, *p* < .001, as well as group membership, with participants preferring Group A members, *B* = 0.48, SE = 0.05, Wald *z* = 9.58, *p* < .001. Again, the interaction was not significant, *B* = 0.01, SE = 0.23 Wald *z* = 0.05, *p* = .964. The random slopes model produced the same pattern (Reward rates: *B* = 2.01, SE = 0.31, Wald *z* = 6.5, *p* < .001, Group bias: *B* = 0.52, SE = 0.19, Wald *z* = 2.74, *p* = .006)

As in Studies 1 and 2, analysis of training phase data produced the same pattern: significant effects of players' reward rate, *B* = 0.93, *SE* = 0.08, Wald *z* = 11.04, *p* < .001, and of group membership, *B* = 0.46, *SE* = 0.04, Wald *z* = 12.09, *p* < .001, on choice. The interaction was not significant, *B* = 0.04, *SE* = 0.12, Wald *z* = 0.40, *p* = .691.

Subjective rewards. Participants' subjective reward estimates were submitted to a linear regression, with actual player reward rate and player group as predictors. Subjective estimates were significantly predicted by players' actual reward rates, *B* = 31.31, *SE* = 8.49, *t* = 3.69, *p* < .001, indicating participants had some knowledge of the reward contingencies (Figure 2A). Subjective reward rates were also predicted by group membership, *B* = 3.88, *SE* = 1.90, *t*(683) = 2.04, *p* = .042, suggesting a weak effect of group bias on subjective reward in addition to the relatively strong effect effects observed on choice behavior.

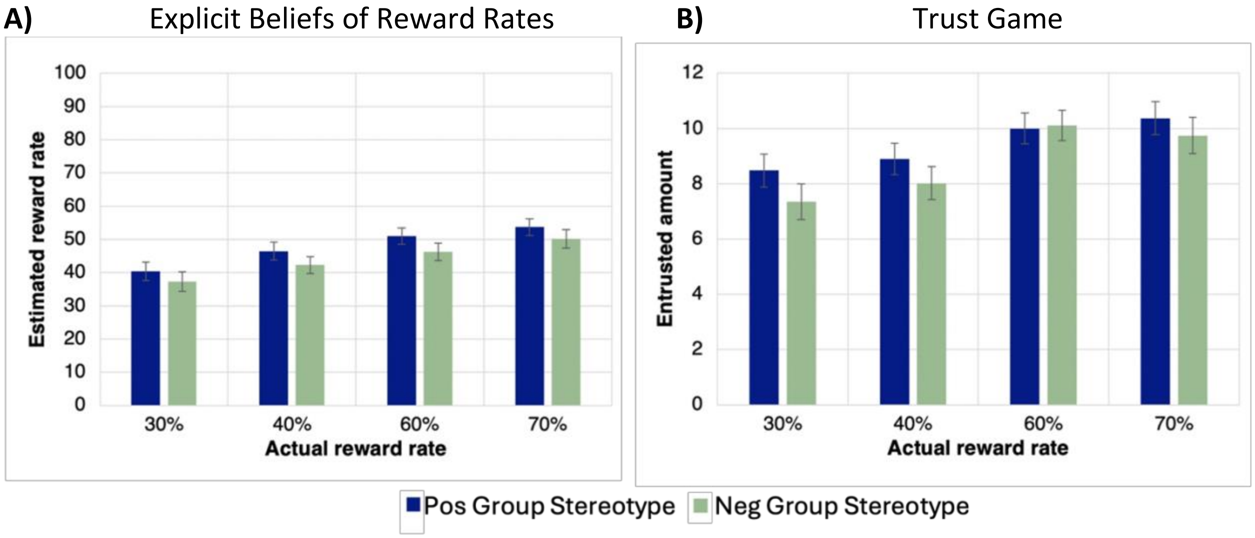

Figure S2. (A) Estimated reward rates for each player. Study 3 participants estimated higher reward rates for more rewarding players and for Group A members. (B) Number of points entrusted to players during the trust game. Study 3 participants entrusted more points to the more rewarding players and to Group A members.

To test whether the effects of group membership on choice behavior could be explained by subjective rewards, or whether the group bias instrumental learning operated independently of subjective rewards, we conducted an analysis in which player reward rates, group membership, and subjective estimates were each included in the main multilevel model predicting test phase choice. We found that while subjective rewards predicted choices to a small extent, *B* = 0.03, *SE* = 0.01, Wald *z* = 26.62, *p* < .001, actual reward rates (*B* = 0.94, *SE* = 0.11, Wald *z* = 5.44, *p* < .001) and player group (*B* = 0.21, *SE* = 0.05, Wald *z* = 3.97, *p* < .001) remained strong predictors of choice behavior. This result suggests that subjective beliefs about group differences in reward did not fully account for the group effect expressed in behavior.

Trust game behavior. Participants' trust game investments were submitted to a linear regression, with players' (trustee) true reward rate and player (trustee) group as predictors. Participants' investments were significantly predicted by players' actual reward rates, *B* = 5.87, *SE* = 1.34, *t*(693) = 4.38, *p* < .001, 95% CI [3.24, 8.50], reflecting that reward learning translated to an expression of trust (Figure 2B). However, the effect of group membership was not significant, *B* = 0.63, *SE* = 0.42, *t*(693) = 1.48, *p* = .140, 95% CI [-0.19, 1.45].

Study 4

Overview. Despite Studies 1-3 consistently showing a transmission of stereotypes to personal preferences, one alternative explanation could be that the stereotype messages provided payoff-relevant information instead of merely creating a generalized positive or negative portrayal. That is, since parts of the stereotypes alluded to differences in wealth levels between the social groups (groups were described as coming from either more or less affluent regions with low or high unemployment), these descriptions may have implicitly communicated information about the expected sharing rates of both groups. In Study 4, we modified the stereotype descriptions used in Studies 1-3 to exclude any wealth-related information.

Method

Participants. Participants were 134 workers on the recruitment platform Connect who received $4.00 for their participation as well as a performance-based monetary bonus, ranging from $0 – $3.00, derived from points earned during the task. We excluded participants who failed to reach a learning criterion of 50% accuracy for 30%-70% player pairs during the test phase, as well as participants who did not finish the main task (N = 29). These exclusions resulted in a final sample size of N = 105 (60 men, 41 women, 1 nonconforming, 3 unreported; M_{age} = 39.50 years, SD_{age} = 12.36 years).

Ethics approval was obtained from the human subjects institutional review board at the University of Amsterdam.

Procedure. The procedure was equivalent to that of previous studies, but the stereotype messages now did not include wealth-related cues. The new stereotype messages were:

Members of Group A live in a secure region, where crime is low and which is commonly seen as peaceful. People from Group A are often perceived to be trustworthy, honest and polite. Members of Group B, in contrast, live in a different region which is considered more dangerous, with high rates of serious crimes such as robbery, assault, and murder. People from Group B are often perceived as hostile, untrustworthy, and dishonest.

Results

Choice behavior. Our analytical approach followed that of previous studies. Multilevel regression predicting player choice again produced significant effects of players' actual reward rate, *B* = 2.29, SE = 0.14, Wald *z* = 16.76, *p* < .001, as well as group membership, with participants preferring Group A members, *B* = 0.36, *SE* = 0.04, Wald *z* = 8.46, *p* < .001 (Figure S3). The random slopes model did not show a significant group effect, *B* = 0.49, SE = 0.31, Wald *z* = 1.57, *p* = .12.

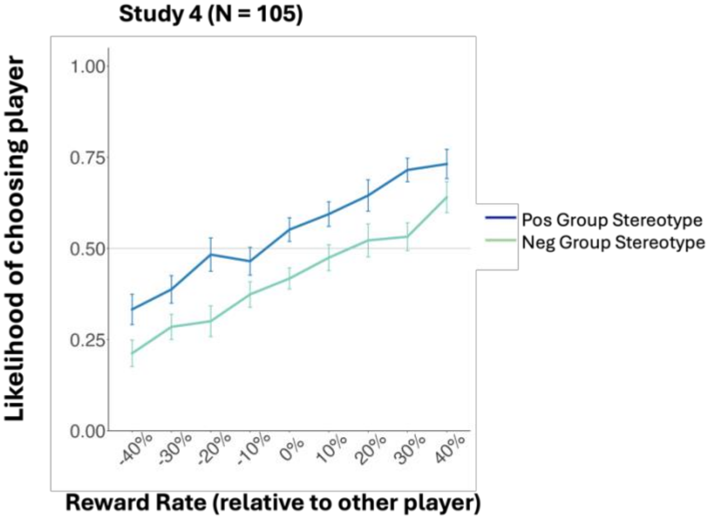

Figure S3. Choice behavior for the test phase of Study 4. Participants' choices (solid lines) demonstrated learning of reward contingencies as well as a group bias. Reward rate, displayed on the x axis, represents the actual reward rate of a given player minus the actual reward rate of the alternative player in a trial. Error bars indicate standard error.

Explicit rewards. A regression with participants' explicit reward ratings and actual reward rates, as well as group membership shows effects for both factors: Reported rewards tracked actual rewards, *B* = 43.44, *SE* = 4.22, *t* = 10.29, *p* < .001, but were also influenced by group, *B* = 4.44, *SE* = 1.34, *t* = 3.32, *p* < .001. Adjusting for reported rewards again retains the group effect in the main analysis, *B* = 0.43, *SE* = 0.03, *t* = 16.6, *p* < .001.

Study 5

Overview. In Study 5, we asked whether the prejudice could be propagated through mere observation of biased choices without any knowledge of group descriptions or direct feedback from group members. In this study procedure, participants observed the training data of a past participant, viewing their choice and feedback but not being exposed to any group descriptions, and then completed their own test phase choices. Each Study 5 participant was yoked to real Study 2 participant ("demonstrator") and observed the training phase behavior of the participant to whom they were yoked. To test whether participants formed a group bias based on this observational learning, Study 5 participants then made their own choices in the test phase. After the test phase of the learning task, participants reported their estimates of player reward rates in the same explicit belief measure as Study 3, followed by estimates of the choice behavior of the demonstrators that they observed. Study 5 was preregistered at https://aspredicted.org/blind.php?x=6zi6fz.

Method

Participants. Participants were recruited from the online NYU subject pool and received course credit for their participation as well as the chance to win a performance-based monetary bonus of $15 to be awarded to the five top performers. Our stopping rule was to collect data until two Study 5 participants were yoked to each of the 62 demonstrators from Study 2 (for a total N of 124). We excluded participants who responded without variation in the post-learning task (*N* = 6), participants who failed to reach 50% on either attention check measure (*N* = 13; see Procedure), and participants who failed to reach a learning criterion of 50% accuracy for the 30%-70% pairs during the test phase (i.e., A-B and G-H; *N* = 33). Because we could not know the number of eventual exclusions during the period of data collection, data were collected from an additional 36 participants who met inclusion criteria but were ultimately not needed and thus excluded. These extra participants were excluded based on chronological order of completing the experiment, and their data were never analyzed. After all exclusions, this approach yielded the target final sample size of *N* = 124 (82 women, 39 men, 3 unreported; M_{age} = 19.5 years, SD_{age} = 1.35 years).

Ethics approval was obtained from the human subjects institutional review board at New York University.

Procedure. Participants read instructions similar to Studies 1 – 3, which explained the nature of the sharing task and the fact that players represented two different groups, but they received no descriptions of the groups. They were also told that they were to learn about the target players by watching past "demonstrators" make decisions and receive feedback. Participants then completed a categorization task of the group membership of target players, which served to both reinforce group membership cues and provide an attention check on which to base participant exclusions. Unlike in Study 3, this categorization task did not include stereotype words; participants only categorized faces of group members. Participants with less than 50% accuracy on the classification task were excluded (*N* = 9).

Next, participants observed the training phase, where, instead of making choices, participants observed the trials of the Study 2 demonstrator to whom they were yoked. Trials were presented in the same order and each yoked trial was animated in real time (using actual reaction times for each prior demonstrator choice) to show choices and subsequent reward feedback, identical to how direct learners viewed choices and feedback. Participants observed the entirety of the yoked training phase, complete with a break between the two blocks, with the exception that trials were skipped if they had originally been excluded in Study 2 based on reaction time. To ensure participants paid attention, "catch" trials appeared after some trials, prompting participants to indicate what choice they had just observed on the previous trial. Twenty catch trials appeared in the observational training phase, occurring in a fixed, pseudorandom order. Participants with less than 50% accuracy on the catch trials were excluded (N = 4).

Participants then completed the test phase, making their own choices, as in Studies 1 – 4. Next, participants reported their explicit beliefs about player reward rates, as in Study 3, typing their response in a box under displays of each player ("How many times out of a hundred would this player share with you?"). Finally, participants reported their estimates of demonstrators' tendency to choose each player ("How many times out of a hundred did the Decider choose this player?").

Results

Observational learning effects. Our analytical approach followed thar of Studies 1 – 4, with a focus on test phase choices. In this study, however, participants did not directly complete a training phase, but instead observed training phase behavior of Study 2 participants. As in the previous studies, in which learning occurred directly, multilevel regression indicated that observational learning produced a significant effect of actual reward rates, B = 1.49, SE = 0.09, Wald z = 16.73, p < .001, as well as a significant effect of group membership, with Study 5 participants preferring Group A players, B = 0.32, SE = 0.04, Wald z = 8.03, p < .001. The interaction was not significant, B = -0.24, SE = 0.18 Wald z = -1.36, p = .173.

As in the preceding experiments, the random effects analysis produced a qualitatively identical pattern (Reward rates: *B* = 1.78, SE = 0.22, Wald *z* = 7.99, *p* < .001, Group bias: *B* = 0.37, SE = 0.19, Wald *z* = 2.0, *p* = .045). This group bias remained after adjusting for Study 2 participants' subjective estimates of player reward rates (*B* = 0.26, *SE* = 0.04, Wald *z* = 6.52, *p* < .001), consistent with an implicit transmission of bias.

It should be noted that this analysis deviated slightly from our preregistered plan, which was to yoke Study 5 participants to the total Study 2 sample (*N* = 78), with two Study 5 participants yoked to each Study 2 participant in order to increase power. This pre-registration did not consider that some Study 2 participants would provide invalid or incomplete date. Hence, in order to obtain validity and rigor, Study 5 participants were yoked only to Study 2 participants included in the final Study 2 analysis. Nevertheless, results were nearly identical using this sample (*N* = 156): test phase learning effect: *B* = 1.49, *SE* = 0.11, Wald *z* = 13.36, *p* < .001; group effect: *B* = 0.39, *SE* = 0.03, Wald *z* = 11.25, *p* < .001; interaction: *B* = -0.17, *SE* = 0.18 Wald *z* = -1.13, *p* = .258.

In addition to these analyses, we also estimated the direct transmission of bias by predicting the participants preference for Group A from the demonstrator´s Group A bias using multilevel regression (*B* = 0.28, SE = 0.09, Wald *z* = 3.21, *p* = .001). This result indicates that the degree of demonstrator group preference was significantly correlated was with the degree of observer group preference.

Explicit beliefs. In the reward estimation task, participants' estimates of targets players' reward rates was not significantly associated with those players' actual reward rates, *B* = -2.48, *SE* = 4.99, *t*(989) = -0.50, *p* = .619, indicating participants had very poor, if any, declarative knowledge of the reward contingencies. There was also no evidence of a group bias in estimations of reward rates, *B* = 0.79, *SE* = 1.58, *t*(107) = 0.50, *p* = .615. Thus, the observational learning of bias appeared to emerge in the absence of explicit beliefs or knowledge regarding player reward rates.

As in Study 3, to more directly test whether choice behaviors reflected a group bias in the absence of explicit beliefs, we tested the main regression with actual player reward rate, group membership, and explicit belief estimates as predictors. Results indicated a small-effect size association between explicit beliefs and choice behavior, *B* = 0.01, *SE* = 0.001, Wald *z* = 20.22, *p* < .001, significant effects remained for actual reward rates, *B* = 1.69, *SE* = 0.13, Wald *z* = 13.09, *p* < .001, and group membership, *B* = 0.26, *SE* = 0.04, Wald *z* = 6.52, *p* < .001.

An analysis of observers' estimate of demonstrator choices indicated that these did not significantly reflect the actual player reward rates, *B* = -0.66, *SE* = 4.77, *t*(989) = -0.14, *p* = .890, or group membership, *B* = 1.05 *SE* = 1.51, *t*(989) = 0.70, *p* = .486.

Study 6

Overview. Previous studies showed that stereotype messages influenced subsequent learning from interactions. To conduct a stronger test of whether stereotype messages, once encoded in memory but not necessarily endorsed, will influence instrumental learning, stereotypes in Study 6 were presented in a context which casts doubts on their veracity. Do stereotype messages affect recipients even if their validity is questioned directly? To answer this question, in Study 6 participants again received group stereotypes, but were subsequently informed that these are merely common stereotype messages which may or may not be true, that individuals within groups vary, and that as a consequence participants should attend to individual player feedback instead of group stereotypes.

Method

Participants. Participants were 148 workers on the recruitment platform Connect who received $4.00 for their participation as well as a performance-based monetary bonus, ranging from $0 – $3.00, derived from points earned during the task. We excluded participants who failed to reach a learning criterion of 50% accuracy for 30%-70% player pairs during the test phase, as well as participants who did not finish the main task (*N* = 42). These exclusions resulted in a final sample size of *N* = 105 (50 men, 44 women, 2 nonconforming, 10 unreported; M_{age} = 37.46 years, SD_{age} = 11.17 years).

Ethics approval was obtained from the human subjects institutional review board at the University of Amsterdam.

Procedure. The procedure was equivalent to that of Study 4, but now the stereotypes appeared together with messages questioning their veracity and usefulness:

"In this study, we are interested in how background information, about people's social groups and where they come from, informs our judgments. Although these groups will be kept anonymous, below are descriptions of how each group is typically viewed:

Members of Group A live in a region typically viewed as secure and with low crime, and it is commonly seen as a peaceful place. People from Group A are often perceived to be trustworthy, honest, and polite.

Members of Group B, by comparison, live in a region considered more dangerous, viewed as having high rates of serious crimes such as robbery, assault, and murder. People from Group B are often perceived as hostile, untrustworthy, and dishonest.

As you see, these two groups are perceived to differ in many ways. However, these descriptions are common stereotypes about these groups and may not be true. Individuals *in each group vary, too, and so the stereotypes can often be misleading. It's important that you learn about the individual players as you engage in repeated interactions in the task."*

Results

Choice behavior. Our analytical approach followed that of previous studies. Multilevel regression predicting player choice again produced significant effects of players' actual reward rate, *B* = 2.33, SE = 0.13, Wald *z* = 17.50, *p* < .001, as well as group membership, with participants preferring Group A members, *B* = 0.57, *SE* = 0.04, Wald *z* = 13.65, *p* < .001 (Figure S4). The random slopes model showed a significant group effect, *B* = 0.86, SE = 0.29, Wald *z* = 2.95, *p* = .0032.

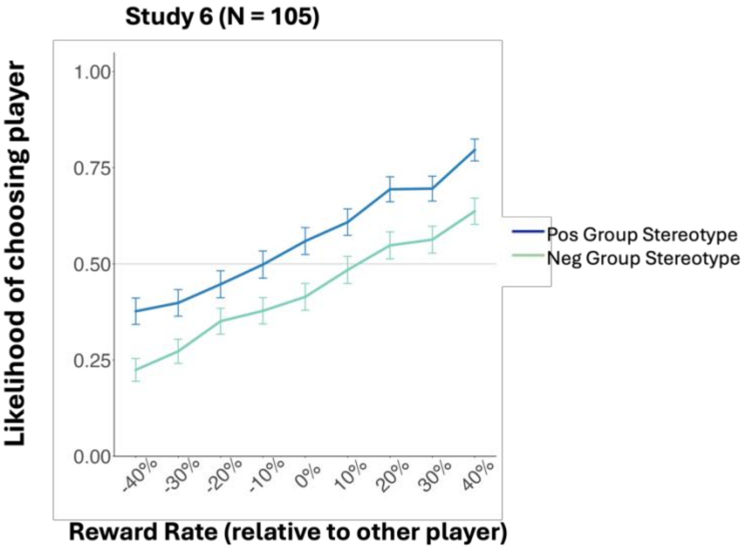

Figure S4. Choice behavior for the test phase of Study 6. Participants' choices (solid lines) were predicted by rewards and target group membership. The x-axis represents the difference between the reward rates of the two available players on a given trial.

Explicit Rewards. Again, reported sharing rates were predicted by group membership, *B* = 4.83, *SE* = 1.41, *t* = 3.42, *p* < .001, and actual reward rates, *B* = 46.05, *SE* = 4.46, *t* = 10.33, *p* <

.001. As in previous studies, the group effect remained significant when adjusting for explicit perception of rewards, $B = 0.48$, $SE = 0.03$, $t = 18.70$, $p < .001$.

Study 7

Overview. The purpose of Study 7 was to replicate the results obtained in Study 6, with one change: To ensure that participants understood that acting on the group stereotypes will reduce their earnings, participants had to pass an understanding quiz after completing the instructions and before starting the learning phase. Only if they correctly indicated that their explicit goal of maximizing rewards required focusing on individual players instead of stereotypes could they proceed to the task.

Method

Participants. Participants were 232 workers on the recruitment platform Connect who received $4.00 for their participation as well as a performance-based monetary bonus, ranging from $0 – $3.00, derived from points earned during the task. We excluded participants who failed to reach a learning criterion of 50% accuracy for 30%-70% player pairs during the test phase, as well as participants who did not finish the main task ($N = 78$). These exclusions resulted in a final sample size of $N = 154$ (87 men, 55 women, 1 nonconforming, 2 other, 9 unreported; $M_{age} = 39.12$ years, $SD_{age} = 11.24$ years).

Ethics approval was obtained from the human subjects institutional review board at the University of Amsterdam.

Procedure. Except for the novel understanding quiz, the procedure was equivalent to that of Study 6. The wording of the understanding quiz was as follows:

"Before you begin the main task, please answer this question about the task instructions: To earn the most points in this task, I should base my choices on
-Descriptions of the player's group
-Feedback from individual players"

Results

Choice behavior. The quiz was correctly completed on the first attempt by 98 participants, by 53 participants on the second attempt, and by 3 participants on the third attempt. Our analytical approach followed that of previous studies: Multilevel regression predicting player choice showed significant effects of players' actual reward rate, *B* = 2.33, SE = 0.11, Wald *z* = 20.90, *p* < .001, as well as group membership, with participants preferring Group A members, *B* = 0.44, *SE* = 0.04, Wald *z* = 12.50, *p* < .001. The random slopes model produced a significant group effect as well, *B* = 0.55, SE = 0.22, Wald *z* = 2.51, *p* = .0012.

Explicit rewards. Reported rewards were predicted by group membership, *B* = 1.98, *SE* = 0.18, *t* = 11.20, *p* < .001, as well as reward rates, *B* = 41.02, *SE* = 0.55, *t* = 75.1, *p* < .001. Including explicit beliefs in the main analysis is consistent with previous results: Group membership predicts choices above and beyond subjective reports, *B* = 0.47, *SE* = 0.02, *t* = 26.00, *p* < .001.

Study 8

Overview. In Study 8, we tested whether a transmission of bias would occur between demonstrators from Study 7 and novel participants. In other words, we aimed to investigate whether prejudice would spread even when demonstrators explicitly tried to avoid the stereotype's influence. As in Study 5, novel participants observed Study 7 participants' behavior and subsequently made their own decisions.

Method

Participants. Participants were 300 workers on Connect who received $4.00 for their participation as well as a performance-based monetary bonus, ranging from $0 – $3.00, derived from points earned during the task. We excluded participants who failed to reach a learning criterion of 50% accuracy for 30%-70% player pairs during the test phase, as well as participants who did not finish the main task (N = 88). We also excluded participants who failed either of two attention measures, as in Study 5 (N = 88). These exclusions resulted in a final sample size of N = 154 (74 men, 72 women, 1 nonconforming, 2 other, 5 unreported; M_{age} = 35.72 years, SD_{age} = 10.53 years).

Ethics approval was obtained from the human subjects institutional review board at the University of Amsterdam.

Procedure. The procedure was equivalent to that of Study 5.

Results

Choice behavior. Again, multilevel regression predicting player choice showed significant effects of players' actual reward rate, B = 1.50, SE = 0.11, Wald z = 13.96, p < .001, as well as group membership, B = 0.19, SE = 0.03, Wald z = 5.67, p < .001. The random slopes model did not produce a significant group effect, B = 0.26, SE = 0.27, Wald z = 0.98, p = .32. Furthermore, observer group bias was predicted by their respective demonstrator bias, B = 0.15, SE =

0.05, *t* = 2.30, *p* = .003. As in previous studies, adjusting for explicit rewards preserved the group effect, *B* = 0.11, SE = 0.16, Wald *z* = 6.89, *p* < .001.

Computational modeling

Our computational modeling analysis evaluated different hypotheses about the mechanisms underlying the group bias observed in the experiments. To this end, we adapted reinforcement learning (RL) and Bayesian learning models previously developed for understanding the influence on verbal instruction on learning (3).

Reinforcement learning. The basis for all reinforcement (RL) models was the standard Q-learning (or Rescorla-Wagner) learning rule:

$$Q_i{}^{t+1} = Q_i{}^{t} + \alpha(R^t - Q_i{}^{t}) \qquad [1]$$

where Q_i is the action value of selecting option *i* in trial *t*, *R* is the reinforcement [no reward = 0, reward = 1] received in trial *t*, and α ($0 \leq \alpha \leq 1$) is a learning rate parameter, which determines how much the difference between the received and the predicted reinforcement (the prediction error) affects subsequent value estimates (4).

In all RL models, the Q-values were transformed into decision probabilities using a standard Softmax function

$$P_i = \frac{e^{Q_i/\beta}}{\sum_{j=1}^{2} e^{Q_j/\beta}} \qquad [2]$$

where *β* (0.01 < *β* ≤ 100) is the temperature parameter that determines the sensitivity of choices to the difference in Q-values. Very low values of *β* results in selecting the action with higher Q-value with probability ~1, while high values of *β* result in explorative choices that are insensitive to the difference in Q-values. Together, equations 1-2 gives an unbiased standard learning model (model 1).

We considered two main mechanisms for group-based bias in RL. First, the semantic information provided in the manipulated group descriptions could result in different *priors*, or initial expectancies, about the value of selecting each group at the outset of the training phase. We implemented this by estimating a prior parameter, *P* (-100 ≤ *P* ≤ 100), which determined the initial Q value for the groups $Q_{Good}^{t=0} = prior, Q_{Bad}^{t=0} = -prior)$. In models without this parameter, the initial Q-values were set to be equal ($Q_{Good}^{t=0} = Q_{Bad}^{t=0} = 0.5$). Non-zero values of the *P* parameter bias initial choices of the group with *P* > 0. We implemented a model with priors but no reward learning in the *bias prior* model (model 2).

In addition to models with such symmetric priors, we evaluated models with separate priors, in which initial expectancies for either group were allowed to vary independently, introducing an extra parameter but allowing for increased flexibility. Models with symmetric priors provided a better fit to the data and their output will be reported in later sections.

It should be noted that, if the model allows for reward learning, the influence of the *P* parameter decreases exponentially across training trials. In other words, experiential learning can rapidly counteract the initial expectancies. We evaluated this in the *bias prior RL* model (model 3).

Second, the *learning rate*, α, might differ between groups, so that participants update Q-values more (or less) rapidly from interacting with one group than the other (eq. 1). To evaluate biased updating, we either estimated α by group (2 α), or by both group and sign of the prediction error (4 α), based on classic social psychological theories that relate prejudice to differential attention to groups (i.e., ingroup favoritism, 49, and outgroup homogeneity, 50) and differential processing of positive and negative behaviors of ingroup vs. outgroup members (i.e., the ultimate attribution error, 51).

*Bayesian learning***.** We also tested how Bayesian learning models accounted for the data. The main motivation for this approach is that Bayesian priors can have a stronger, more long-lasting effect on behavior than in the RL framework we describe above (where the prior is just the initial Q-value). We used standard Bayesian beta-binomial learning models (3), which explicitly estimate the probability of reward for selecting each group, given a beta distributed prior with hyperparameters α and β (both initialized to 1 for each stimulus *i*). The model learned by updating α and β (for each stimulus) by adding the running count of reward and no-reward feedback (separately for each stimulus *i*). Given a beta prior, this amounts to calculating the posterior distribution for each stimulus using Bayes rule:

$$\alpha_i{}^{t+1} = \alpha_i{}^{t} + pos \qquad [3]$$

$$\beta_i{}^{t+1} = \beta_i{}^{t} + neg \qquad [4]$$

where *pos* = 1 after reward feedback, and 0 after no-reward feedback, and vice-versa for *neg*. In addition, the running counts are decayed multiplicatively on each trial by a free parameter γ

(0 ≤ γ ≤ 1), which allows the model to forget potentially outdated information (8). Choices were probabilistically taken (following a Softmax function, eq. 2) by comparing the modes of the posterior distributions:

$$mode_i = \frac{\alpha_i - 1}{\alpha_i + \beta_i - 2} \qquad [5]$$

We evaluated two versions of this model. In the first version (model 9), the initial α parameter for Group A was estimated. In this model formulation, both the mode and the precision of the prior is affected by α. More evidence is required to counteract a precise prior. If α_{Good} is higher than a α_{Bad}, the model is biased to select group A. The second model (model 10) incorporated an additional parameter w (1 ≤ w ≤ 100), which modulated the feedback in a manner congruent with the semantic information (i.e., a confirmation bias). For stimuli from Group A, this gives

$$\alpha_i^{\,t+1} = \alpha_i^{\,t} + wpos \qquad [6]$$

$$\beta_i^{\,t+1} = \beta_i^{\,t} + \frac{1}{w} neg \qquad [7]$$

In other words, the model learns faster from positive outcomes and slower from negative outcomes. For stimuli from Group B, the effect of w was the inverse (i.e., faster learning from negative outcomes and slower learning from positive outcomes).

We exploratorily evaluated two Bayesian models for Studies 1-3, primarily to allow for a stronger influence of the prior. As they provided poor fit to the data, we did not evaluate these models in later studies.

Model #	**Conceptual label**	**parameters**	**# parameters**
1	*Unbiased learning*	α,β	2
2	*Stereotype-only*	β, P	2
3	*Stereotype-individuation*	α,β,P	3
4	*Group-learning*	$\alpha_{Good},\alpha_{Bad},\beta$	3
5	*Stereotype-learning*	$\alpha_{Good},\alpha_{Bad}$, P,β	4
6	*gain/loss group-learning*	$\alpha_{Good}+,\alpha_{Good}-$, $\alpha_{Bad}+$, $\alpha_{Bad}-,\beta$	5
7	*Stereotype-gain/loss group learning*	$\alpha_{Good}+,\alpha_{Good}-$, $\alpha_{Bad}+,\alpha_{Bad}-,P$, β	6

Table S1. Overview of tested models.

Parameter estimation. Parameter estimation was conducted using the maximum-likelihood approach, which finds the set of parameters that maximize the probability of the participant´s trial-by-trial test phase choices given the model. Optimization was done by to minimizing the negative log-likelihood, *-L*, computed by:

$$-L = -\sum_{t=1}^{T} ln\ (P_{choice}(t)) \qquad [8]$$

where *T* denotes the total number of trials. Parameters were independently fitted to the test phase data for each participant using the Nelder-Mead optimization method. To avoid local minima in parameter fitting, optimization was initiated with 60 randomly selected start values. Model implementations and parameter fitting was done in *R* 3.5.1.

Model comparison. Model comparison was primarily based on the Akaike Information Criterion (AIC), a measure of goodness of fit of a model that penalizes complexity (9):

$$AIC = -2\ln(L) + 2k \qquad [9]$$

where *–ln(L)* is the negative log-likelihood and *k* is the number of model parameters. A smaller AIC hence indicates a better model fit.

Model comparison was based on the sum AIC across participants. For simplicity, we present the *ΔAIC*, which is the difference between model *i* and the best fitting model.

Table S2 shows the ΔAIC for each experiment separately, as well as the combined ΔAIC. Model 5, combining biased prior expectations and biased learning rates for each group, fit the data best in experiments 1-2, 4, and 6-7. Experiment 3 was best fit by model 7, which included separate learning rates for positive and negative prediction errors (equation 1) from group A and B. However, the difference between model 5 and 7 in Experiment 3 was relatively small. To formally test for the reliability of the apparent difference between experiments, we used a random-effects approach. Specifically, we used a linear mixed model with AIC as the dependent variable, and participant as random factor to test the interaction between experiment and model. This approach showed a main effect of model, *F*(7, 3420) = 2.79, p = .004, indicating that model 5 had significantly lower AIC than the other models. However, there interaction between model and experiment was not reliable, *F*(24, 3420) = 0.83, p = .69, indicating that model 5 provided the best fit across experiments. In addition, Bayesian model comparison indicated that the posterior probability that all four experiments had the same model frequency was *P* = 0.96. Combining all experiments, we also find that the exceedance probability of model 5 was the most common among the candidate models was 1. Together, these results indicate that a combination of biased priors and biased learning rates best accounted for the influence of group descriptive information on instrumental learning across experiments.

Model	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Experiment 1	511	1140	213	175	0	252	228
Experiment 2	709	703	474	71	0	843	1141
Experiment 3	884	1536	196	199	70	114	0
Experiment 4	1397	2004	201	304	0	374	154
Experiment 6	743	1833	86	203	0	216	193
Experiment 7	1975	3392	640	692	0	677	306

Table S2. ΔAIC by model. The table shows the difference in AIC, summed across participants, for each model relative to the best fitting model (with ΔAIC = 0**).**

Relation between model parameters and group-based bias. To understand in more detail how the parameters of Model 5 related to the stereotype bias observed in choice behavior, we regressed the estimated model parameters (excluding the Softmax temperature *β*) onto the degree of choice preference for positively-stereotyped group members in the test phase (proportion of *Group A* choices). All model parameters were rank transformed and standardized to improve linearity and interpretability. We conducted this analysis for Studies 1-3. We found that both the prior *P* (*β* = 0.084, SE = 0.012, t = 6.62, p < .0001) and the learning rate parameter for the negatively-stereotyped group ("Group B"; α_{bad}: *β* = -0.039, SE = 0.013, t = -3.7, p = .003) statistically predicted the degree of group preferences. In other words, a larger initial value difference between the groups (the prior), together with a lower learning rate for Group B, was associated with a stronger bias. The learning rate for the positively-stereotyped group (“Group A”) was not reliably related to bias (α_{good}: β = -0.003, SE = 0.011, t = -0.33, p = .74).

Supporting Information References

1. P. G. Devine, A. J. Elliot, Are racial stereotypes really fading? The Princeton trilogy revisited. Pers. Soc. Psychol. Bull. 21, 1139–1150 (1995).

2. A. P. Gregg, B. Seibt, M. R. Banaji, Easier done than undone: asymmetry in the malleability of implicit preferences. J. Pers. Soc. Psychol. 90, 1–20 (2006).

3. B. B. Doll, W. J. Jacobs, A. G. Sanfey, M. J. Frank, Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. Brain Res. 1299, 74–94 (2009).

4. R. A. Rescorla, A. R. Wagner, "A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement" in Classical Conditioning II: Current Research and Theory, A. H. Black, W. F. Prokasy, Eds. (Appleton- Century-Crofts, 1972), pp. 64–99.

5. M. B. Brewer, The psychology of prejudice: Ingroup love and outgroup hate? J. Soc. Issues 55, 429–444 (1999).

6. B. Park, M. Rothbart, Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. J. Pers. Soc. Psychol. 42, 1051–1068 (1982).

7. T. F. Pettigrew, The ultimate attribution error: Extending allport's cognitive analysis of prejudice. Pers. Soc. Psychol. Bull. 5, 461–476 (1979).

8. Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. Neuroimage, 46, 1004-1017.

9. N. D. Daw, "Trial-by-trial data analysis using computational models" in Decision Making, Affect, and Learning, (Oxford University Press, 2011), pp. 3–38