Journal of Experimental Psychology: General

ISSN: 0096-3445

2025, Vol. 154, No. 2, 347–357
https://doi.org/10.1037/xge0001694

Costly Exploration Produces Stereotypes With Dimensions of Warmth and Competence

Xuechunzi Bai[1], Thomas L. Griffiths[2, 3], and Susan T. Fiske[2, 4]
[1] Department of Psychology, The University of Chicago
[2] Department of Psychology, Princeton University
[3] Department of Computer Science, Princeton University
[4] Department of Psychology, School of Public and International Affairs, Princeton University

Traditional explanations for stereotypes assume that they result from deficits in humans (ingroup-favoring motives, cognitive biases) or their environments (majority advantages, real group differences). An alternative explanation recently proposed that stereotypes can emerge when exploration is costly. Even optimal decision makers in an ideal environment can inadvertently form incorrect impressions from arbitrary encounters. However, all these existing theories essentially describe shortcuts that fail to explain the multidimensionality of stereotypes. Stereotypes of social groups have a canonical multidimensional structure, organized along dimensions of warmth and competence. We show that these dimensions and the associated stereotypes can result from *feature-based* exploration: When individuals make self-interested decisions based on past experiences in an environment where exploring new options carries an implicit cost and when these options share similar attributes, they are more likely to separate groups along multiple dimensions. We formalize this theory via the contextual multiarmed bandit problem, use the resulting model to generate testable predictions, and evaluate those predictions against human behavior. We evaluate this process in incentivized decisions involving as many as 20 real jobs and successfully recover the classic dimensions of warmth and competence. Further experiments show that intervening on the cost of exploration effectively mitigates bias, further demonstrating that exploration cost per se is the operating variable. Future diversity interventions may consider how to reduce exploration cost, in ways that parallel our manipulations.

Public Significance Statement

Stereotypes are multidimensional, including features that go beyond sheer good–bad valence. Current psychological theories, which focus on social, cognitive, and sample biases, do not explain the origins of such complex stereotypes. In this article, we show that a novel psychological mechanism can reproduce the multidimensional stratification of social groups and the resulting complex stereotypes: When individuals make self-interested decisions based on past experiences in an environment where exploring new options carries an implicit cost and when options share similar attributes, they are more likely to separate groups along multiple dimensions. A further set of intervention experiments provides causal evidence that reducing exploration cost can substantially mitigate even complex stereotypes.

Keywords: stereotype, warmth–competence, explore–exploit, generalization, intervention

Supplemental materials: https://doi.org/10.1037/xge0001694.supp

This article was published Online First November 21, 2024.

Ross Otto served as action editor.

This work has been presented at the following venues, and the authors are grateful for attendees' insightful feedback: Organizational Behavior at the Yale School of Management, Massachusetts Institute of Technology Sloan School, and Stanford Graduate School of Business; Department of Psychology at the University of California Davis, the University of Washington, the University of Chicago, and New York University; Annual Meetings of the Academy of Management at Boston, Society of Personality and Social Psychology at San Diego, and Behavioral Decision Research in Management at Booth School of Business; Computational Social Science workshop at the University of Chicago; and International Conference on Computational Social Science in Philadelphia.

This article was funded by John Templeton Foundation (Grant 23900-G0002-10012712-101) awarded to Susan T. Fiske.

Xuechunzi Bai played a lead role in conceptualization, data curation, formal analysis, investigation, methodology, project administration, validation, visualization, writing–original draft, and writing–review and editing. Thomas L. Griffiths played a lead role in supervision and an equal role in conceptualization, formal analysis, investigation, methodology, project administration, writing–original draft, and writing–review and editing. Susan T. Fiske played a lead role in funding acquisition and an equal role in conceptualization, investigation, methodology, project administration, supervision, writing–original draft, and writing–review and editing.

Correspondence concerning this article should be addressed to Xuechunzi Bai, Department of Psychology, The University of Chicago, Kelly Hall 409, 5848 South University Avenue, Chicago, IL 60637, United States. Email: baix@uchicago.edu

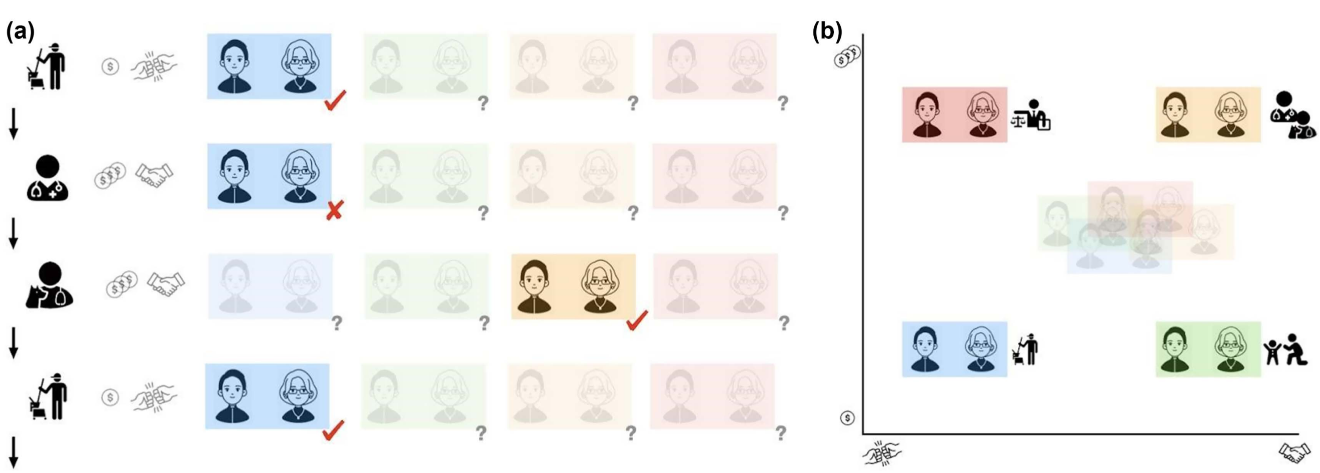
Social stereotypes seem to be a fundamental part of human societies. They organize expectations about gender, race, nationality, and appearance and carry associations about perceived trustworthiness and competence (Bai et al., 2020; Bian et al., 2017; Katz & Braly, 1933; Todorov et al., 2015). People often learn these complex stereotypes from segregated societal structures signaling, for example, social status and cooperative intent (Fiske et al., 2002; Koenig & Eagly, 2014). What position a specific group occupies in such structures depends on complex economic, cultural, historical, and political circumstances. However, the mechanisms that differentiate groups follow basic psychological principles. In this article, we use a combination of computational simulations and incentivized behavioral experiments to show that trade-offs between exploring new options and following past experiences can produce multidimensional traits that recapitulate the axes along which people represent real social groups: Differentiated stereotypes emerge spontaneously when exploration is costly and is guided by socially constructed features.

Existing psychological explanations for social stratification between groups have focused on four causes: biased decision makers, particularly those who are high status and powerful, assigning minorities to disadvantageous positions to protect their ingroup or to oppress outgroups (Altemeyer, 1983; Brewer, 1999; Jost & Banaji, 1994; Pratto et al., 1994); cognitively limited decision makers having distorted mental representations due to inherent constraints such as memory capacity or attention selectivity (Fiske & Taylor, 1984; Hamilton & Gifford, 1976; Macrae et al., 1994; Sherman et al., 2000; Trope & Thompson, 1997); statistically unsophisticated decision makers not taking into account that they are observing unrepresentative samples, producing biases (Denrell, 2005; Fiedler, 2000; Payne et al., 2017); and, most controversially, actual group differences resulting in groups being sorted into different positions (Eagly & Steffen, 1984; McCauley et al., 1995). These four explanations thus attribute the origins of stereotypes to a defect in human decision making or in the environmental samples.

Contrary to these notions, a recently proposed fifth perspective, informed by work in computer science, highlights the inherent trade-off between "exploring" new options and "exploiting" existing knowledge (Sutton & Barto, 2018) and posits that even optimal decision makers might inadvertently produce bias when exploring unfamiliar options entails an implicit cost (Bai et al., 2022). While these five accounts might explain why people differentiate between groups, particularly identifying an ingroup as good and an outgroup as less good, they do not explain more complex stereotypes that go beyond a simple good–bad dichotomy (Abele et al., 2021; Koch et al., 2016; Nicolas et al., 2022; Zou & Cheryan, 2017). For example, stereotypes of immigrants in the United States are not merely binary; perceptions vary in a multifaceted manner: Russians are seen as competent but untrustworthy, Mexicans are perceived as neither competent nor trustworthy, Native Americans are seen as friendly but not competent, and Canadians are perceived as capable and friendly (Bai et al., 2020; Lee & Fiske 2006). We build on the explore–exploit framework to show that multidimensional social stratification need not to be rooted in flaws in humans or the environments. It is simply a consequence of the way social decisions are often posed. Costly exploration, combined with socially constructed features that provide a basis for generalization, is sufficient to produce rich multidimensional stereotypes.

To illustrate our proposed mechanism and to anticipate the methods used in our experiments, imagine a manager hiring individuals from different social groups for different jobs (Figure 1). The manager's goal is to ensure successful outcomes in these jobs. Assume that individuals from all groups are equally and highly likely to succeed in all kinds of jobs. The manager does not know this and seeks to learn how well the different groups perform based

Figure 1
The Hiring Task as a Contextual Multiarmed Bandit



Note. An example illustrates how making new decisions based on past (selective) experiences can create a stratified unit that produces multidimensional stereotypes that are incorrect. Panel (a) shows example jobs, their associated features such as social status and cooperative intent, and four groups. Each decision only has one group being hired, whose performance is then revealed and is used to guide new decisions, while the other three groups remain unknown. Panel (b) shows mental representations after these decisions are made. The example mental map is organized by two features—competence and trustworthiness. The true situation is pictured in the background while the incorrect impressions of the groups formed by the decision maker are shown in the foreground. All colors denote group membership. See the online article for the color version of this figure.

on experience. Unfortunately, the learning process suffers from a serious constraint: The manager can only observe the performance of people they hire, so they remain ignorant of how well the people they did not hire could have done.

As a specific example, consider five jobs that vary on two features: high-status and high-trust doctors and veterinarians; high-status and low-trust lawyers; low-status and high-trust childcare aides; low-status and low-trust garbage collectors (Fiske & Dupree, 2014; Koenig & Eagly, 2014). As jobs become available one by one, the manager assigns people from different groups to do each job in turn and observes the performance of the hired individuals.

Initially, when a garbage collector position opens, the manager may randomly choose a person from one group (blue) without enough information to make a better decision. But they learn that it is a good choice. Next, the manager must choose somebody for a doctor position, still without enough evidence to support a definitive decision, so perhaps they want to stick to the same group one more time but quickly discover it is a poor decision (suppose they happen to hit the rare incompetent individual in this population where most groups can do most jobs). A third job, a veterinarian position, is available. Although the manager has not hired a veterinarian before, given that veterinarians share similar features with doctors, managers may generalize from their past experiences. Given their past negative experience with blues as doctors, the manager may switch to a different group (yellow). They learn that the newly recommended individual performs well. The process continues.

Remember, the underlying probability of being successful is identical and high for all pairs of jobs and groups. Despite individual variation, on average, every group is just as good as any other group at performing all jobs. Intuitively, initial positive experiences recommending members of one group for garbage collectors may encourage the manager to recommend more members from that group as garbage collectors or for similar jobs. Consequently, the manager is less likely to recommend people from other groups for the same positions or people from that group for other jobs. If so, the manager has introduced social stratification, hiring more people from one group for low-status and low-trust jobs. Observing this pattern, the manager and others might wrongly conclude that the overrepresented group in these positions is incompetent and untrustworthy.

This example illustrates how a series of seemingly adaptive decisions can produce a social reality that sorts members of different groups into distinct positions, without needing to appeal to group motives, cognitive limits, sample imbalances, or group differences. This behavior is adaptive for the individual decision maker as it optimizes hiring performance in two key ways. First, it minimizes the implicit cost from exploring a new uncertain group, which might not perform as reliably as a more familiar choice (Bai et al., 2022). Second, it further reduces the exploration cost by generalizing shared features across positions. Using these features, the decision maker can recommend similar but not identical positions to the same group (Shepard, 1987). Despite multiple adaptive benefits to the individual, this behavior is detrimental to society because the byproduct of these decisions is a biased and stratified representation of reality. Not only do some groups receive inadequate exploration, but the underlying features associated with them also become the foundation for complex and multidimensional stereotypes. Multidimensional stratification emerges from adaptive individual decisions for the individual, but decisions that are maladaptive for the collective.

This minimal explanation for the origin of stereotypes is challenging to test because multiple mechanisms are confounded in studies of stereotypes based on real-world knowledge. To address this challenge, we used a combination of computational modeling and incentivized behavioral experiments. The computational model precisely defines the problem being solved and demonstrates the emergence of stereotypes in the absence of group motivations, cognitive limitations, unequal sample size, or differing group qualities (see the Model section below). The behavioral experiment enriches the simple scenario assumed in the model with as many as 20 real-world jobs. Both computational agents and human participants stratify their environments and form stereotypes, even along multiple dimensions, simply because feature-based exploration has intrinsic costs (see the Experiment section below). Intervening to reduce these costs, however, reduces stratification and stereotypes (see the Evaluating Interventions section below).

Method

Materials

Experimental details, data set construction, analysis details, formal modeling, and computational simulations are provided in the Supplemental Material.

Transparency and Openness

We report power analyses, preregistration reports, study materials, simulation codes, analysis codes, and anonymized raw data for all reported studies on the Open Science Framework, which can be accessed by everyone at https://osf.io/6p8wu/?view_only=22709d2fd3164a90880af4b0e2679f7f.

Results

Model

To formalize our hiring problem, we adapt the contextual multiarmed bandit task—a fundamental problem explored in theoretical treatments of sequential decision making and reinforcement learning in computer science and related disciplines (Sutton & Barto, 2018). In a multiarmed bandit task, an agent chooses actions (pulling an "arm" of the "bandit," an old-fashioned gambling machine) to receive rewards over multiple rounds. Each arm has a probability distribution over rewards. In each round, the agent selects an arm and receives a reward sampled with the corresponding probability. The agent wants to maximize their cumulative rewards but is unaware of the reward distributions associated with the arms. The agent thus needs to balance two competing options: *exploring* a new arm to learn its reward and *exploiting* the arm that is known to give the highest expected reward.

Many real decisions involve choosing between options that are differentiated by observable features. The *contextual* multiarmed bandit task captures this by assuming that the reward distribution depends not only on the arm but also on a set of features that describe the decision context on that round (L. Li et al., 2010). Instead of estimating the reward distribution for each arm, the agent now estimates the function that maps contextual features to reward distributions. While this problem is harder to solve than the simple multiarmed bandit, it yields greater flexibility as the agent can learn to generalize to future similar but not identical situations based on their features. This is the critical modification that makes multidimensional stereotypes emerge.

While there are no known optimal solutions for the contextual bandit task, we use a Bayesian approach called Thompson sampling (Agrawal & Goyal, 2012; Thompson, 1933). Thompson sampling uses Bayesian inference to estimate the probability of reward associated with each arm and then samples an arm with a probability that matches the posterior probability of that arm offering the best chance of reward. This approach has been shown to be an effective model of human choices and social interactions (Bai et al., 2022; Schulz et al., 2018). To learn the function between contextual features and reward distributions, we employ Bayesian logistic regression (Chapelle & Li, 2011; L. Li et al., 2010).

Using the described model, we simulated the behavior of adaptive-decision agents who follow Thompson sampling and random-decision agents who do not maximize rewards or use past experiences in choosing among four groups over 40 choice trials (see Supplemental Material for model details). The choices involved allocating members of the different groups to jobs, where each job had a known set of features reflecting the need for trustworthiness and competence, and the adaptive-decision agents' estimated parameters for each group indicating the extent to which they had these features. The underlying rate at which rewards were delivered to all groups was the same: Rewards were sampled from a Bernoulli distribution where each individual had a 90% chance of succeeding in the job, hence, delivering a reward for the decision maker. Note that the simulated agents are initialized with an uninformative prior that follows a unit normal distribution for each group. The agents do not have parameters for group motivation or memory limitations, and the ground truth data set does not contain unequal population sizes or different reward probabilities (see other simulation variants such as differing ground truth and differing prior beliefs in Supplemental Material).

Nonetheless, the simulation reveals that adaptive-decision agents, while attempting to maximize rewards through past experience, are more likely to allocate groups differentially and form stereotypes compared to random-decision agents. We illustrate this using an ordinary least squares linear regression model with the agent type as the predictor variable (adaptive coded as 1 vs. random coded as 0) and the entropy of the distribution of choices over groups (i.e., choice entropy) and the distance between estimated parameters for the groups (i.e., stereotype dispersion) as the outcome variables. This model shows that the adaptive-decision agents show a lower entropy, indicative of stratified choices ($b = -.645$, 95% confidence interval, CI [−.614, −.676], $p < .001$), and a bigger distance, indicative of differentiated stereotypes ($b = 1.447$, 95% CI [1.596, 1.297], $p < .001$; Figure 2 for prototypes) as compared to the random-decision agents. Stratified choices and dispersed estimated parameters emerge from the agents trying to solve the explore–exploit dilemma to maximize their rewards while minimizing the hidden cost of exploring the unknown.

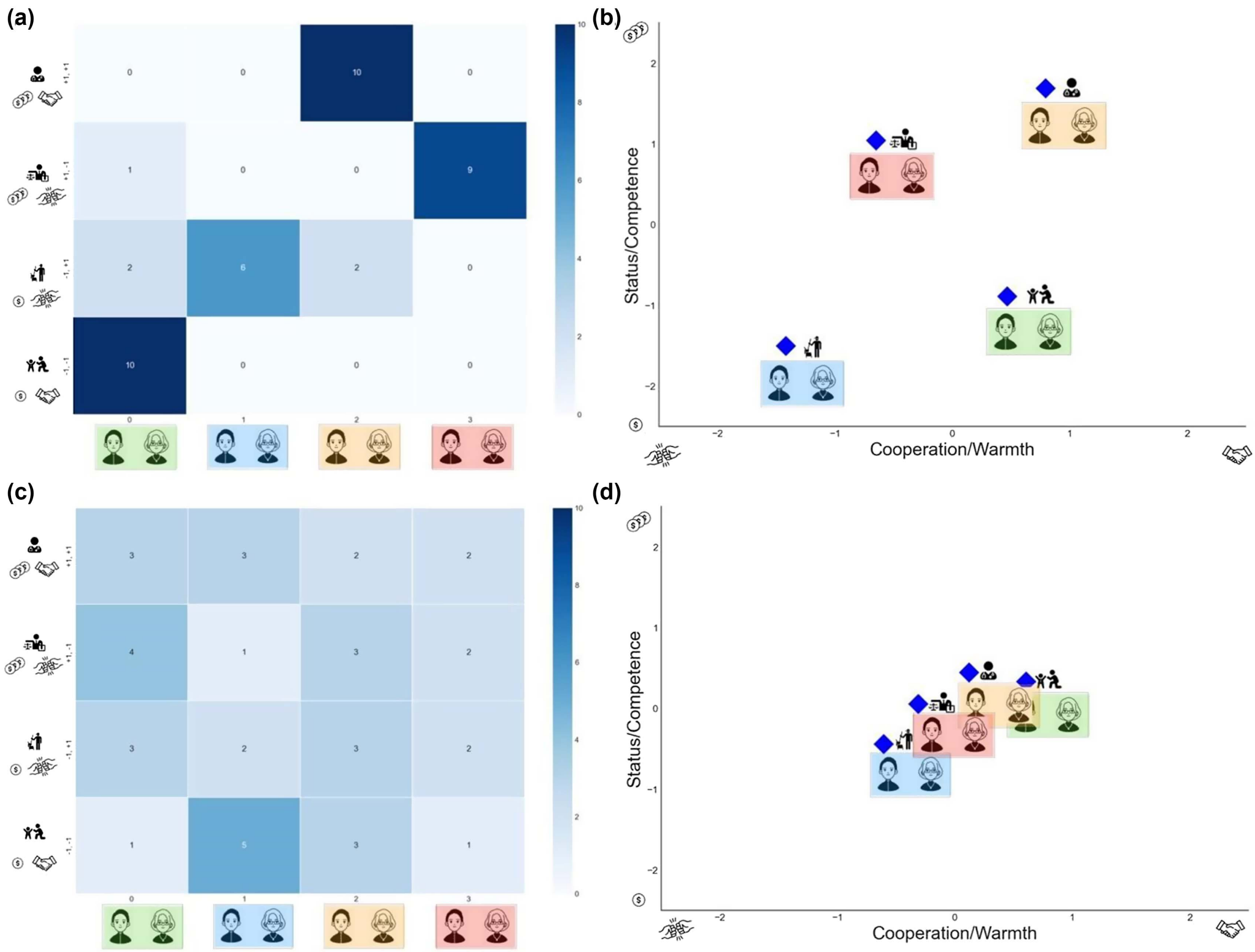
Experiment

We tested the predictions of this model in a large-scale online experiment in which participants (total $N = 1,310$) made hiring decisions involving novel social groups. Participants were told that they had been recruited by the mayor of a made-up place, Toma City, to recommend members of four groups of people, the Tufas, Aimas, Rekus, and Wekis, for different jobs. The better recommendations the participants make, the more money they earn. To test whether participants generalize their experiences from a few limited jobs to a large amount of similar but not identical jobs, we prepared 20 different kinds of jobs (Fiske & Dupree 2014; Koenig & Eagly, 2014; see Supplemental Material for a preliminary study in norming these jobs), and jobs open one at a time at random. In the adaptive exploration condition, participants make decisions sequentially and learn the outcome of their recommendation after each decision, earning 1 point or 0 points. In the random exploration condition, participants observe the mayor making random decisions. This minimal design aimed to reduce the impact of group motivations, cognitive limitations, unrepresentative sampling, and quality differences while focusing on the causal effects of adaptive versus random exploration (see Supplemental Material for experimental designs). Specifically, we controlled for group motivations by using novel unfamiliar groups that do not belong to any participants; we controlled for selective attention by designing a small number of trials with straightforward presentation to minimize cognitive load; we controlled for differences in group size and true reward probabilities by fixing the parameters to be identical across groups. To test our central claim about the role of adaptive exploration, we did include one key comparison: demonstrating that "lesioning" adaptive exploration removed the effect.

Participants were online workers from the Cloud Research high-quality subject pool who speak English as their first language and are older than 18 years old. Self-report demographic shows the average age was 40; 51% female, 46% male, 1% nonbinary; 74% White, 10% Black, 6% Hispanic, 5% Asian, 4% multiracial; 75% hold some college or bachelor's degree; the average political orientation was slightly liberal with an average score of 3.94 on a scale from 1 = *extremely conservative* to 6 = *extremely liberal*. These demographics reflect typical characteristics of online American workers for psychological studies (see Supplemental Material for more details).

Confirming the model predictions, the human data show statistically significant differences in choice entropy between the adaptive exploration condition and the random exploration condition ($b = -.476$, 95% CI [−.437, −.514], $p < .001$). This analysis controls for individual differences in age, gender, race, education, and political orientation. Participants who make their own decisions display lower entropy, corresponding to more stratified and unequally distributed choices (Figure 3a "Default"; see an example participant in Figure 4a). In contrast, participants who observe random decisions from the mayor display higher entropy with less stratified and more equally distributed choices (Figure 3a "Ideal"; see an example participant in Figure 4c). Moreover, compared to participants who observe random decisions, participants who adaptively explore are more likely to report larger mental distances in the trustworthiness–competence space ($b = .343$, 95% CI [.597, .089], $p < .001$; Figure 3b Default vs. Ideal; see example participants in Figure 4b and 4d, respectively). The stratified choice also holds for imagined future hires where participants make new decisions regarding unseen applicants. The stereotype dispersion also holds for status and cooperation dimensions, which are theorized as structural antecedents of competence and trustworthiness (Abele et al., 2021; Fiske et al., 2002; see Supplemental Material for more results).

Figure 2
Two Example Simulated Results



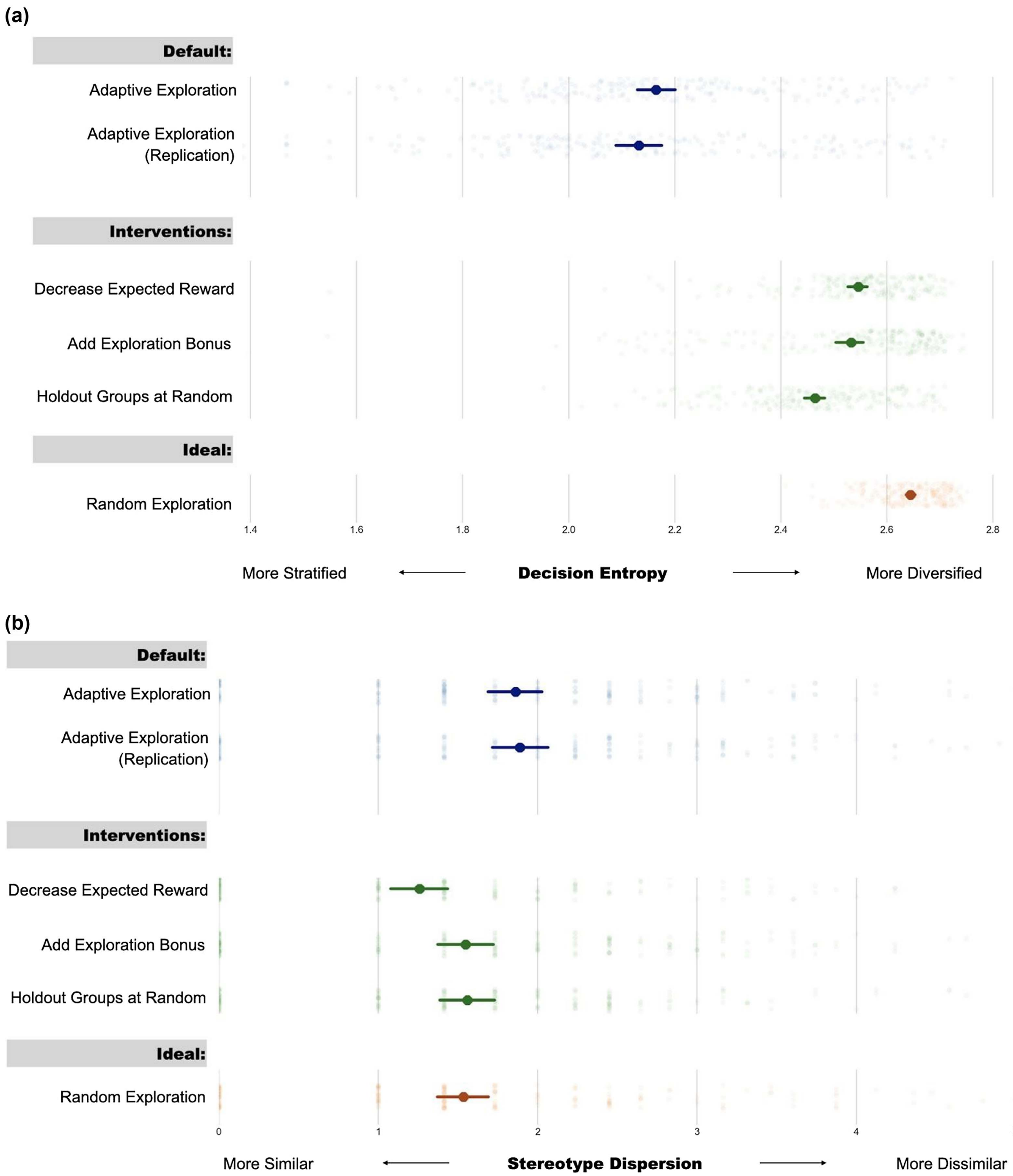
Note. In Panels (a) and (b): From an agent who makes adaptive decisions. In Panels (c) and (d): From another agent who makes decisions at random. The heatmaps on the left panels show how many times a group, on the horizontal axis, is recommended for a job, on the vertical axis. The scatterplots on the right panels show estimated coefficients for the four groups on the two binary features. For aggregate simulation results see Supplemental Material simulation section. All colors denote group membership. See the online article for the color version of this figure.

Two results are worth highlighting: First, we see evidence for the emergence of multidimensional stereotypes. As shown in Figure 4b, participants do not simply polarize Toma groups as the uniformly good versus the utterly bad ones. Rather, they clearly differentiate along at least two dimensions—for example, Tufas are competent but not trustworthy or Wekis are incompetent but trustworthy (Figure 4a). Second, we see evidence for generalization (Shepard, 1987). Regardless of the diversity of the jobs, participants clearly find (dis)similarities between jobs. As shown in Figure 4a, participants do not randomly assign jobs to people, but rather, they cluster jobs into reasonable categories and use the generalized category to guide decisions. For example, once participants discover Rekus are good custodians, they then assign Rekus to be cashiers and dishwashers even though they never have direct experience of Reku cashiers or Reku dishwashers because they perceive custodians as similar to cashiers and dishwashers. These two results highlight the unique contribution of this work, which is how feature-based exploration enables the emergence of multidimensional stereotypes. In sum, human behavioral data replicate the model predictions, showing that a stratified society emerges from participants acting adaptively to solve the explore–exploit trade-off and that this stratification leads to multidimensional stereotypes with a similar structure to those observed-for-real social groups.

Evaluating Interventions

If the implicit cost of exploration is the key mechanism that results in multidimensional stratification, intervening on this cost should reduce stratification and stereotypes. We studied three interventions to test this prediction: adding an exploration bonus, decreasing the reward probability, and imposing a random holdout.

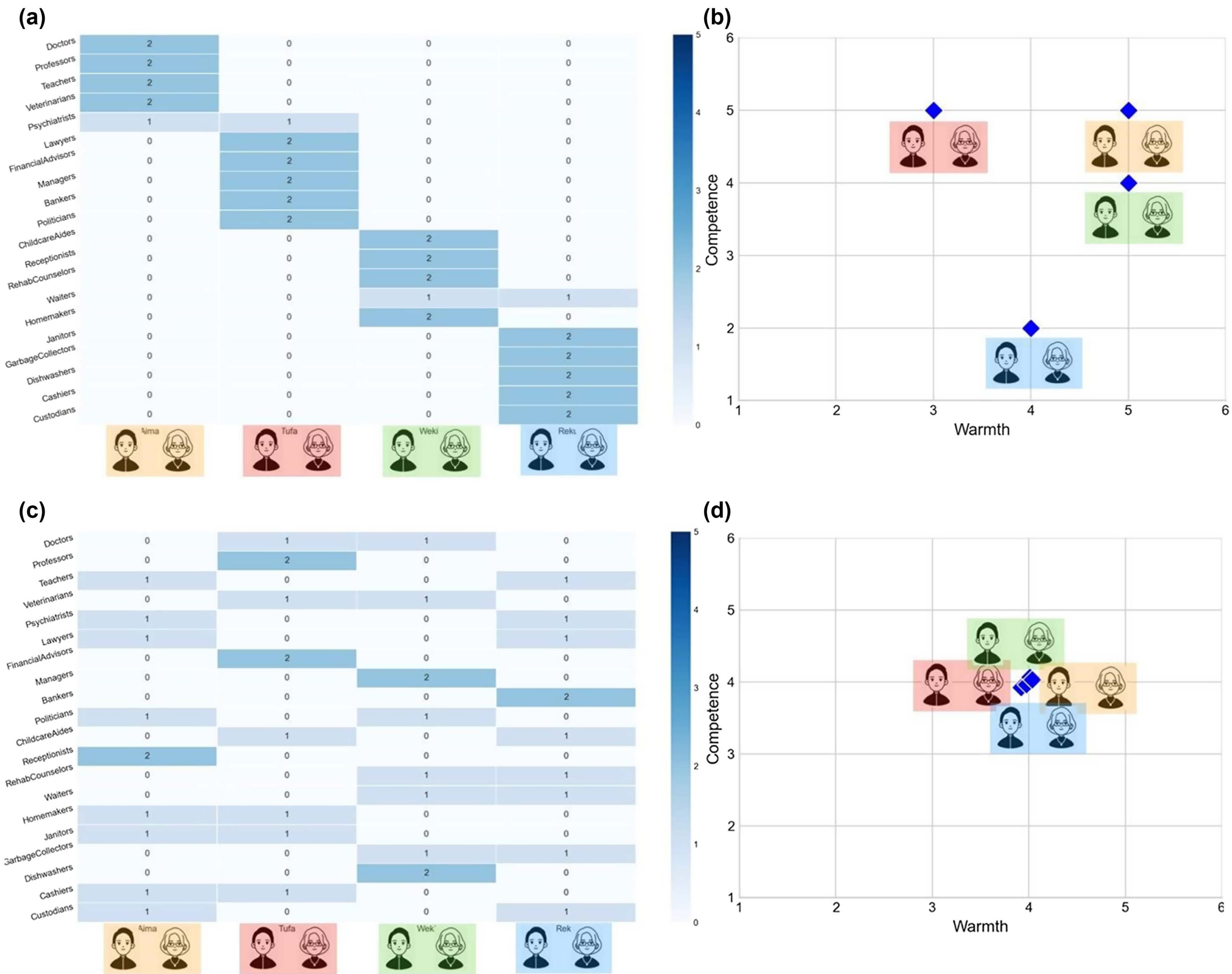
Figure 3
Average Treatment Effects in Human Behavioral Experiments



Note. The vertical axis represents experimental conditions: The default panel with blue bars shows the adaptive exploration condition where participants make their own hiring decisions in the main study and replication in the mechanism study. The ideal panel with orange bars shows the random exploration condition where participants observe the mayor making random decisions. The intervention panel with green bars shows three interventions that manipulate the exploration cost to diversify choices and reduce stereotypes. The horizontal axis represents the average treatment effects for hiring choices in terms of choice entropy in Panel (a) and stereotype dispersion in Panel (b). Panel (a) shows more stratified choices to more diversified choices in the order of the default exploration, the interventions, and the random ideal condition. Panel (b) shows more dissimilar to similar stereotypes in the order of the default exploration, the interventions, and the random ideal condition. In all graphs, error bars represent bootstrapped 95% confidence intervals. See the online article for the color version of this figure.

Figure 4

Prototypes of Stratified Versus Diversified Hiring Choices and Dissimilar Versus Similar Stereotypes



Note. Panels (a) and (b) show results from Participant No. 153, who was assigned to the adaptive exploration condition. (a) This participant predominantly selects Aimas to work in high-status, high-trust jobs; Tufas in high-status, low-trust jobs; Wekis in low-status, high-trust jobs; and Rekus in low-status, low-trust jobs. (b) As a result of such stratified choices, this participant thinks Aimas are warm (trustworthy) and competent, Tufas are competent but not warm, Wekis are incompetent but warm, and Rekus are neither competent nor warm. Panels (c) and (d) show results from Participant No. 281, who was assigned to the random exploration condition. (c) This participant observes the mayor selecting randomly. (d) As a result, this participant thinks Aimas, Tufas, Wekis, and Rekus are similarly warm and competent. Note that not every participant has the same strong pattern; there is considerable individual heterogeneity (see Supplemental Material for details). All colors denote group membership. See the online article for the color version of this figure.

Each intervention addresses the implicit cost of exploration in a different way. First, adding a bonus to untried options directly incentivizes exploration (Bellemare et al., 2016). This design uses a common approach adopted in reinforcement learning to encourage exploration, which counts how many times a state has been encountered and assigns a bonus to the state that has rarely been visited (Bellemare et al., 2016). Second, decreasing the reward probability to make all groups less likely to yield rewards should make it less likely that people quickly encounter a successful group, meaning that they need to explore more. This design originates from empirical observations that environments with scarce rewards foster more exploration than environments with abundant rewards (Bai et al., 2022; Harris et al., 2020). Third, randomly holding out some groups to make them unavailable forces exploration, making the cost of exploration irrelevant. This design is inspired by natural events such as travel restrictions due to the pandemic. By restricting which groups are available for exploration, people may try groups that they would not have chosen based on past experiences, thus, increasing the chances of trying novel options.

We initially tested these interventions using our computational model, which showed that all three interventions resulted in more diverse choices and more similarity among the estimated parameters of the groups (see Supplemental Material for detailed modeling results). Briefly, an increase in choice entropy and decrease in stereotype dispersion is found to be a function of the unit price of exploration bonus (Supplemental Figure S9), the expectation of the chance of getting a reward (Supplemental Figure S10), and the likelihood that two groups are unavailable when the agents need to make a decision (Supplemental Figure S11).

Guided by the simulations, we then tested these interventions in a behavioral experiment. Human participants were randomly assigned to one of the four conditions (N = 807): The control condition proceeds with the same hiring scenario as the adaptive exploration condition of our original experiment; the exploration bonus condition adds a diversity bonus, and it displays the sum of rewards from hiring decisions throughout the experiment; the lower reward condition decreases the underlying reward probabilities without an explicit change in instructions; the random holdout condition adds a travel restriction that randomly affects different groups, making two groups unclickable most of the time (see Supplemental Material for experimental designs).

Participants were online workers from Connect, the high-quality online platform hosted by Cloud Research, who speak English as their first language and are older than 18 years old. Self-report demographics show that the average age was 40; 50% female, 50% male; 67% White, 11% Black, 9% Asian, 6% Hispanic, and 4% multiracial; 71% of participants hold some college or bachelor's degree; the average political orientation was slightly liberal with an average score of 3.98 on a scale from 1 = *extremely conservative* to 6 = *extremely liberal*. The average score for this task was 4.7 out of 5, indicating acceptable engagement among participants (details in Supplemental Material).

Consistent with the model, participants made more exploratory hiring when they were assigned to the exploration bonus (b = .390, 95% CI [.340, .440], p < .001), lower reward (b = .402, 95% CI [.355, .449], p < .001), and random holdout (b = .319, 95% CI [.272, .366], p < .001) conditions than those in the control condition (Figure 3a "Interventions" and "Default replication"). There are consistent, although weaker, treatment effects on the distances between the estimated parameters of the four Toma groups. Compared to the baseline, participants reveal smaller distances on the trustworthiness–competence space in the exploration bonus (b = −.339, 95% CI [−.603, −.074], p = .012), lower reward (b = −.693, 95% CI [−.959, −.427], p < .001), and random holdout (b = −.294, 95% CI [−.557, −.30], p = .029; Figure 3b "Interventions" and "Default replication") conditions. This pattern is robust for future hires and status cooperation dimensions (see Supplemental Material for more results). Interventions that change the cost of exploration are thus promising avenues for mitigating stratification and stereotypes.

Discussion

The mechanism of feature-based exploration that we have introduced in this article makes several innovative contributions. First, it provides a plausible explanation for the emergence of multidimensional stereotypes rather than those based purely on valence. Without assuming deficits in either decision makers or the environments, feature-based exploration explains how multidimensional stratification and stereotypes can emerge when decision makers need to minimize exploration cost by both exploiting past experiences and generalizing from limited experiences to similar but not identical contexts. In an incentivized hiring experiment, using as many as 20 diverse real jobs, this mechanism is sufficient to reproduce the warmth-by-competence space that people use to represent real social groups. Learning that one group is good at doing one category of jobs and using that experience to guide category-sensitive decisions is adaptive for the individual because it minimizes exploration costs. Nonetheless, this strategy brings collateral damage to society because it leaves other groups underexplored for certain types of jobs, resulting in stratification along dimensions that guide interpersonal interactions. Second, our intervention studies are the first to show that exploration cost per se is the operative variable. Introducing bonus rewards for diverse hires, assessing candidates using challenging tasks, and randomly making some groups unavailable for selection effectively reduces the cost of exploration, diversifies decisions, and reduces stereotypes in our artificially constructed online hiring experiment.

Our proposed mechanism complements but differs from prior theories on the origin of stereotypes, as follows. (1) The motivation to maximize self-interest can be orthogonal to the motivation to maintain group identity or hierarchy (e.g., Brewer, 1999). Identifying the causes of stratification and stereotypes as pursuing self-interest with exploration yields very different interventions. Complementing strategies such as creating a common ingroup identity (Gaertner & Dovidio, 2009), our proposal suggests changes in the reward structure for exploration. Consistent with the call for structural changes to redress social bias, our mechanism provides concrete ideas such as introducing bonus rewards for diverse hires. (2) A lack of exploration differs from confirmation bias or metacognitive myopia (e.g., Hamilton & Gifford, 1976). To see why, disentangle two different goals. The incentive in our task is to maximize rewards (earn as many points as possible) whereas the incentive in confirmation bias and metacognitive myopia is to strengthen beliefs (learn the underlying principles as accurately as possible). Although it has been assumed that to maximize rewards one needs to maximize accuracy, we show that the two goals do not always align. Hence, inaccuracy can arise not as a cognitive limitation, but as a side-effect of trying to maximize rewards (see also Le Mens & Denrell, 2011; Rich & Gureckis, 2018). (3) Our proposed mechanism does not depend on asymmetric population sizes when one group is more accessible than other groups (e.g., Alves et al., 2018). Adding unbalanced population size may exacerbate this effect; however, one should not forget that the definitions of majority and minority are not fixed either. Rather than starting with a fixed majority/minority representation, our mechanism provides a process that may create such asymmetry: Individuals who are not explored enough become the numerical minority. (4) Our proposed mechanism does not endorse stereotype accuracy (e.g., Jussim, 2017) because we showed that inaccurate stereotypes emerge even when the ground truth is otherwise.

Most importantly, none of the above theories demonstrably explains why stereotypes have more than one dimension. In contrast, we find the diverse contents of stereotypes associated with social groups could be a result of generalization based on socially constructed features of different jobs. Given that identical situations are rarely encountered twice, the ability to generalize is a crucial adaptive mechanism for humans (Schulz et al., 2018; Shepard, 1987). However, when this generalization process is coupled with decisions to balance exploration and exploitation, it can lead to wrongful association of certain features with specific groups. Absent evidence from less explored alternatives, people might consistently apply these generalized features in future judgments, laying the ground for multidimensional stereotypes. If jobs or social roles were restricted to a single valence dimension, we would expect to see stereotypes represented merely by positivity and negativity. Yet, our empirical evidence—a large sample of ecologically valid jobs—indicates that human participants perceive jobs varying across at least two dimensions, supporting the plausibility of multidimensional stereotype framework.

The goal of using the contextual multiarmed bandit model is to demonstrate that the factor we focus on—adaptive decisions based on dimensions of the environment—can result in stereotype formation. This demonstration enables us to express more precise and quantifiable hypotheses that guide the design of our human experiments. Our critical results are the tests of those qualitative hypotheses to show how our theoretical explanation can result in stereotype formation, not whether the quantitative predictions of the model align with human performance, as we are not intending it as a model about explicit cognitive mechanism. Other mechanisms are plausible but not necessary, as our study causally manipulates the role of experience-based exploration and demonstrates that lesioning experience-based exploration removes the effect. One limitation of this current model is that this model does not account for all responses provided by every single participant in our experiments. There is considerable individual heterogeneity; some participants' decisions are more similar to those of the simulated Bayesian agents while others show more discrepancy (see Supplemental Material for details on individual heterogeneity). Future work can study the interaction between exploration, individual differences, and other mechanisms to deliver a more comprehensive understanding of the origin of stereotypes.

Social scientists have studied diversity and stereotypes from either an individual or a structural lens. However, the new mechanism we have identified suggests that the culprit may be an interaction of the two. It challenges the common assumption that unjust systems are either the result of prejudiced or cognitively stressed decision makers or the result of power maintaining or undiversified organizational arrangements. Instead, it highlights the possibility that unjust systems can also be created by locally adaptive, reward-maximizing decision makers. A company merely pursuing its profit can hire certain groups of workers for specialized tasks but underexplore other groups for inexperienced tasks (D. Li et al., 2020). A university merely pursuing a higher ranking for research can admit certain kinds of researchers for particular disciplines but underexplore other combinations (Wapman et al., 2022). These reasonable local decisions in the short term can create stratified broader societal structures in the long term.

Some real-world policy implications of this idea range well beyond employment discrimination. For example, one pertains to refugee resettlement. Policymakers and social scientists, leveraging large-scale data sets and machine-learning algorithms, propose allocating refugees with similar demographic features to specific locations for similar jobs based on past success (Bansak et al., 2016). Such a plan can be suitable for refugees in the short term because it brings more satisfaction and contributes to the local economy. However, this plan, our model predicts, will cause future damage in the form of multidimensional stereotyping and data-driven discrimination.

The exploration cost mechanism that produces stereotypes in humans also provides a psychological analog of fairness concerns in artificial intelligence. For instance, recommendation algorithms often attempt to infer user preferences based on their past behaviors. However, these algorithms may inadvertently limit exposure to diverse options, making some unreachable to users (Dean et al., 2020). While optimizing customer engagement may be an adaptive strategy for the local algorithm, it simultaneously perpetuates stratification in the global online system.

Constraints on Generality

Although our results show that stereotypes can result from feature-based exploration in a simplified laboratory setting, we caution against broader real-world generalizations. First, our participants may not be representative. Although our recruited human participants constitute large and standard samples for psychological studies (see Supplemental Material for demographic details), they may not represent those who make hiring decisions in the real world. Hence, this proposed mechanism needs to be tested with participants such as managers in organizations, administrators in educational institutions, or officers in immigration offices. Second, some interventions are not easily transferable. For instance, the exploration bonus approach may face institutional backlash. It requires the institution to change the incentive structure of their hiring objectives to explicitly reward demographic diversity with money. Likewise, randomly holding out some groups to make resources unavailable to them can cause ethical concerns. However, a quasi-experimental approach can help. Using natural experiments, such as the introduction of travel restrictions due to the pandemic, researchers can leverage policy changes as the exogenous shock to assess whether and how exploratory behaviors and stereotypes have changed.

Conclusion

Stereotypes are shared cultural beliefs, and segregation is a collective endeavor. Future work should study how idiosyncratic and biased individual experiences become entrenched, not mitigated, within collective systems (Lyons & Kashima, 2003; Martin et al., 2014). Our approach extracts the minimal conditions under which stereotypes can emerge, but it needs real-world corroboration. Future work can use historical, immigration, or organizational data sets to examine adaptive exploration in everyday choices (Card et al., 2022; Charlesworth et al., 2022). Costly exploration should be added to the list of psychological mechanisms that can lead to stereotypes, creating an opportunity for future research that integrates these different mechanisms (Almaatouq et al., 2024). However, continuing to ignore the role of exploration in the creation of stereotypes will reinforce the very injustices that we seek to eradicate. Scientists and practitioners should design systems that facilitate exploration in social decision making, and the interventions explored in this article provide a first step in that direction.

References

Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review*, *128*(2), 290–314. https://doi.org/10.1037/rev0000262

Agrawal, S., & Goyal, N. (2012, June). Analysis of Thompson sampling for the multi-armed bandit problem. *Conference on learning theory* (pp. 39.1–39.26). JMLR Workshop and Conference Proceedings.

Almaatouq, A., Griffiths, T. L., Suchow, J. W., Whiting, M. E., Evans, J., & Watts, D. J. (2024). Beyond playing 20 questions with nature: Integrative experiment design in the social and behavioral sciences. *Behavioral and Brain Sciences*, *47*, Article e33. https://doi.org/10.1017/S0140525X22002874

Altemeyer, B. (1983). *Right-wing authoritarianism*. University of Manitoba Press. https://doi.org/10.1515/9780887558894

Alves, H., Koch, A., & Unkelbach, C. (2018). A cognitive-ecological explanation of intergroup biases. *Psychological Science*, *29*(7), 1126–1133. https://doi.org/10.1177/0956797618756862

Bai, X., Fiske, S. T., & Griffiths, T. L. (2022). Globally inaccurate stereotypes can result from locally adaptive exploration. *Psychological Science*, *33*(5), 671–684. https://doi.org/10.1177/09567976211045929

Bai, X., Ramos, M. R., & Fiske, S. T. (2020). As diversity increases, people paradoxically perceive social groups as more similar. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(23), 12741–12749. https://doi.org/10.1073/pnas.2000333117

Bansak, K., Hainmueller, J., & Hangartner, D. (2016). How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers. *Science*, *354*(6309), 217–222. https://doi.org/10.1126/science.aag2147

Bellemare, M., Srinivasan, S., Ostrovski, G., Schaul, T., Saxton, D., & Munos, R. (2016). *Unifying count-based exploration and intrinsic motivation*. arXiv. https://doi.org/10.48550/arXiv.1606.01868

Bian, L., Leslie, S. J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*(6323), 389–391. https://doi.org/10.1126/science.aah6524

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues*, *55*(3), 429–444. https://doi.org/10.1111/0022-4537.00126

Card, D., Chang, S., Becker, C., Mendelsohn, J., Voigt, R., Boustan, L., Abramitzky, R., & Jurafsky, D. (2022). Computational analysis of 140 years of US political speeches reveals more positive but increasingly polarized framing of immigration. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(31), Article e2120510119. https://doi.org/10.1073/pnas.2120510119

Chapelle, O., & Li, L. (2011). *An empirical evaluation of Thompson sampling* [Conference session]. Proceedings of the 24th International Conference on Neural Information Processing Systems, Granada, Spain.

Charlesworth, T. E. S., Caliskan, A., & Banaji, M. R. (2022). Historical representations of social groups across 200 years of word embeddings from Google Books. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(28), Article e2121798119. https://doi.org/10.1073/pnas.2121798119

Dean, S., Rich, S., & Recht, B. (2020, January). Recommendations and user agency: The reachability of collaboratively filtered information. *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 436–445).

Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review*, *112*(4), 951–978. https://doi.org/10.1037/0033-295X.112.4.951

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*(4), 735–754. https://doi.org/10.1037/0022-3514.46.4.735

Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review*, *107*(4), 659–676. https://doi.org/10.1037/0033-295X.107.4.659

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Fiske, S. T., & Dupree, C. (2014). Gaining trust as well as respect in communicating to motivated audiences about science topics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(Suppl. 4), 13593–13597. https://doi.org/10.1073/pnas.1317505111

Fiske, S. T., & Taylor, S. E. (1984). *Social cognition*. McGraw-Hill.

Gaertner, S. L., & Dovidio, J. F. (2009). A common ingroup identity: A categorization-based approach for reducing intergroup bias. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 489–505). Psychology Press.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, *12*(4), 392–407. https://doi.org/10.1016/S0022-1031(76)80006-6

Harris, C., Fiedler, K., Marien, H., & Custers, R. (2020). Biased preferences through exploitation: How initial biases are consolidated in reward-rich environments. *Journal of Experimental Psychology: General*, *149*(10), 1855–1877. https://doi.org/10.1037/xge0000754

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, *33*(1), 1–27. https://doi.org/10.1111/j.2044-8309.1994.tb01008.x

Jussim, L. (2017). Précis of social perception and social reality: Why accuracy dominates bias and self-fulfilling prophecy. *Behavioral and Brain Sciences*, *40*, Article e1. https://doi.org/10.1017/S0140525X1500062X

Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *The Journal of Abnormal and Social Psychology*, *28*(3), 280–290. https://doi.org/10.1037/h0074049

Koch, A., Imhoff, R., Dotsch, R., Unkelbach, C., & Alves, H. (2016). The ABC of stereotypes about groups: Agency/socioeconomic success, conservative–progressive beliefs, and communion. *Journal of Personality and Social Psychology*, *110*(5), 675–709. https://doi.org/10.1037/pspa0000046

Koenig, A. M., & Eagly, A. H. (2014). Evidence for the social role theory of stereotype content: Observations of groups' roles shape stereotypes. *Journal of Personality and Social Psychology*, *107*(3), 371–392. https://doi.org/10.1037/a0037215

Le Mens, G., & Denrell, J. (2011). Rational learning and information sampling: On the "naivety" assumption in sampling explanations of judgment biases. *Psychological Review*, *118*(2), 379–392. https://doi.org/10.1037/a0023010

Lee, T. L., & Fiske, S. T. (2006). Not an outgroup, not yet an ingroup: Immigrants in the Stereotype Content Model. *International Journal of Intercultural Relations*, *30*(6), 751–768. https://doi.org/10.1016/j.ijintrel.2006.06.005

Li, D., Raymond, L. R., & Bergman, P. (2020). *Hiring as exploration* (No. w27736). National Bureau of Economic Research.

Li, L., Chu, W., Langford, J., & Schapire, R. E. (2010, April). A contextual-bandit approach to personalized news article recommendation. *Proceedings of the 19th international conference on world wide web* (pp. 661–670).

Lyons, A., & Kashima, Y. (2003). How are stereotypes maintained through communication? The influence of stereotype sharedness. *Journal of Personality and Social Psychology*, *85*(6), 989–1005. https://doi.org/10.1037/0022-3514.85.6.989

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, *66*(1), 37–47. https://doi.org/10.1037/0022-3514.66.1.37

Martin, D., Hutchison, J., Slessor, G., Urquhart, J., Cunningham, S. J., & Smith, K. (2014). The spontaneous formation of stereotypes via cumulative cultural evolution. *Psychological Science*, *25*(9), 1777–1786. https://doi.org/10.1177/0956797614541129

McCauley, C. R., Jussim, L. J., & Lee, Y. T. (1995). *Stereotype accuracy: Toward appreciating group differences*. American Psychological Association.

Nicolas, G., Bai, X., & Fiske, S. T. (2022). A spontaneous stereotype content model: Taxonomy, properties, and prediction. *Journal of Personality and Social Psychology*, *123*(6), 1243–1263. https://doi.org/10.1037/pspa0000312

Payne, B. K., Vuletich, H. A., & Lundberg, K. B. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, *28*(4), 233–248. https://doi.org/10.1080/1047840X.2017.1335568

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*(4), 741–763. https://doi.org/10.1037/0022-3514.67.4.741

Rich, A. S., & Gureckis, T. M. (2018). The limits of learning: Exploration, generalization, and the development of learning traps. *Journal of Experimental Psychology: General*, *147*(11), 1553–1570. https://doi.org/10.1037/xge0000466

Schulz, E., Konstantinidis, E., & Speekenbrink, M. (2018). Putting bandits into context: How function learning supports decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(6), 927–943. https://doi.org/10.1037/xlm0000463

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*(4820), 1317–1323. https://doi.org/10.1126/science.3629243

Sherman, J. W., Macrae, C. N., & Bodenhausen, G. V. (2000). Attention and stereotyping: Cognitive constraints on the construction of meaningful social impressions. *European Review of Social Psychology*, *11*(1), 145–175. https://doi.org/10.1080/14792772043000022

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

Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, *25*(3–4), 285–294. https://doi.org/10.1093/biomet/25.3-4.285

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Trope, Y., & Thompson, E. P. (1997). Looking for truth in all the wrong places? Asymmetric search of individuating information about stereotyped group members. *Journal of Personality and Social Psychology*, *73*(2), 229–241. https://doi.org/10.1037/0022-3514.73.2.229

Wapman, K. H., Zhang, S., Clauset, A., & Larremore, D. B. (2022). Quantifying hierarchy and dynamics in US faculty hiring and retention. *Nature*, *610*(7930), 120–127. https://doi.org/10.1038/s41586-022-05222-x

Zou, L. X., & Cheryan, S. (2017). Two axes of subordination: A new model of racial position. *Journal of Personality and Social Psychology*, *112*(5), 696–717. https://doi.org/10.1037/pspa0000080

Received March 9, 2024
Revision received July 31, 2024
Accepted October 4, 2024 ■