Multidimensional Stereotypes Emerge Spontaneously When Exploration is Costly

Abstract

Stereotypes of social groups have a canonical multidimensional structure, reflecting the extent to which groups are considered competent and trustworthy. Traditional explanations for stereotypes – group motives, cognitive biases, minority/majority environments, or real-group differences – assume that they result from deficits in humans or their environments. A recently-proposed alternative explanation – that stereotypes can emerge when exploration is costly – posits that even optimal decision-makers in an ideal environment can inadvertently create incorrect impressions. However, existing theories fail to explain the multidimensionality of stereotypes. We show that multidimensional stratification 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 multi-armed bandit problem, use the resulting model to generate testable predictions, and evaluate those predictions against human behavior. In particular, we evaluate this process in incentivized decisions involving as many as 20 real jobs, and successfully recover the classic warmth-by-competence stereotype space. 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, such as introducing bonus rewards for diverse hires, assessing candidates using challenging tasks, and randomly making some groups unavailable for selection.

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. Here 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.

Main Text

Introduction

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 (*1-4*). People often learn these complex stereotypes from segregated societal structures signaling, for example, social status and cooperative intent (*5-6*). 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. Incorrect impressions of the abilities of different groups can emerge purely because of individuals who make social decisions facing an implicit cost for exploring new options (*7*). Here, we use a combination of computational simulations and incentivized behavioral experiments to show that the same mechanism can produce multidimensional stereotypes 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 in order to protect their ingroup or to oppress outgroups (*8-11*); cognitively limited decision-makers having distorted mental representations due to inherent constraints such as memory capacity or attention selectivity (*12-16*); statistically unsophisticated decision-makers not taking into account that they are observing unrepresentative samples, producing biases (*17-20*); and, most controversially, actual group differences resulting in groups being sorted into different positions (*21-22*). 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 that highlights the inherent tradeoff between "exploring" new options and "exploiting" existing knowledge (*23-25*) posits that even optimal decision-makers might inadvertently produce bias when exploring unfamiliar options entails an implicit cost (*7*). While these five accounts might explain why people differentiate between groups, particularly identifying an in-group as good and an out-group as less good (*26*), they do not explain more complex stereotypes that go beyond a simple good-bad dichotomy. For example, stereotypes of immigrants in the US are not merely binary; perceptions vary in a multifaceted manner: Russians are seen as competent but untrustworthy, Mexicans are neither competent nor trustworthy, Native Americans as friendly but not competent, and Canadians are capable and friendly (*3, 27*). 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 (Fig. 1). The manager's goal is to ensure successful outcomes in these jobs. Assume 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 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 (*5, 28*). 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.

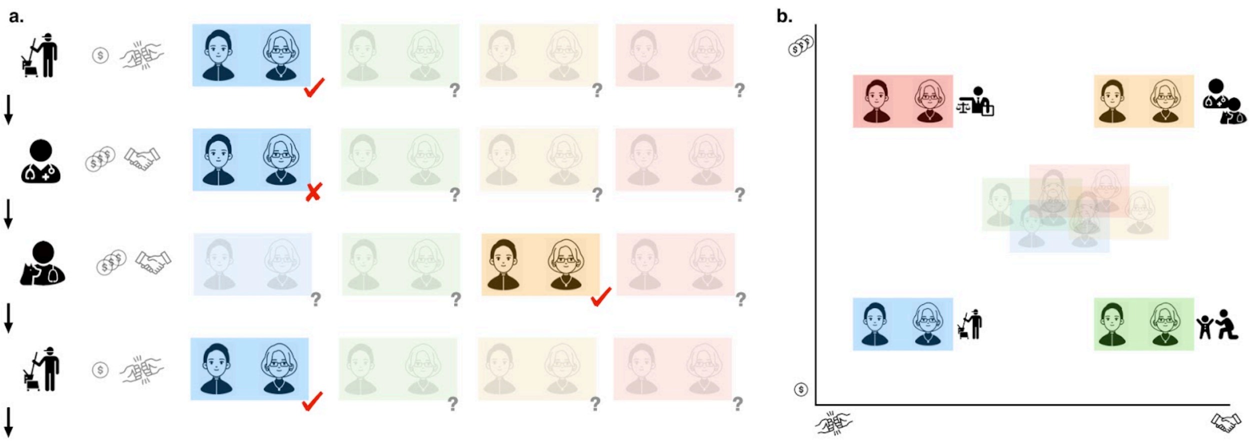

Fig. 1. The hiring task as a contextual multi-armed bandit. An example illustrates how making new decisions based on past (selective) experiences can create a stratified unit that produces multi-dimensional stereotypes that are incorrect. Panel a shows example jobs, their associated features such as social status and cooperative intent, and four candidate 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.

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 (*7*). Second, it further reduces the exploration cost by leveraging shared features across positions. Using these features, the decision-maker can recommend similar but not identical positions to the same group. 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, 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. 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. Intervening to reduce these costs however reduces stratification and stereotypes.

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 (*23*). 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 (*29*). 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 (*30-32*). 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 (*7, 33*). To learn the function between contextual features and reward distributions, we employ Bayesian logistic regression (*29, 32*).

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 *SI* 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 do not have parameters for group motivation or memory limitations, and the ground truth dataset does not contain unequal population sizes or different reward probabilities.

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% CI [-.614, -.676], p < .001), and a bigger distance, indicative of differentiated stereotypes (b = 1.447, 95% CI [1.596, 1.297], p < .001; Fig. 2 for example agents) 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.

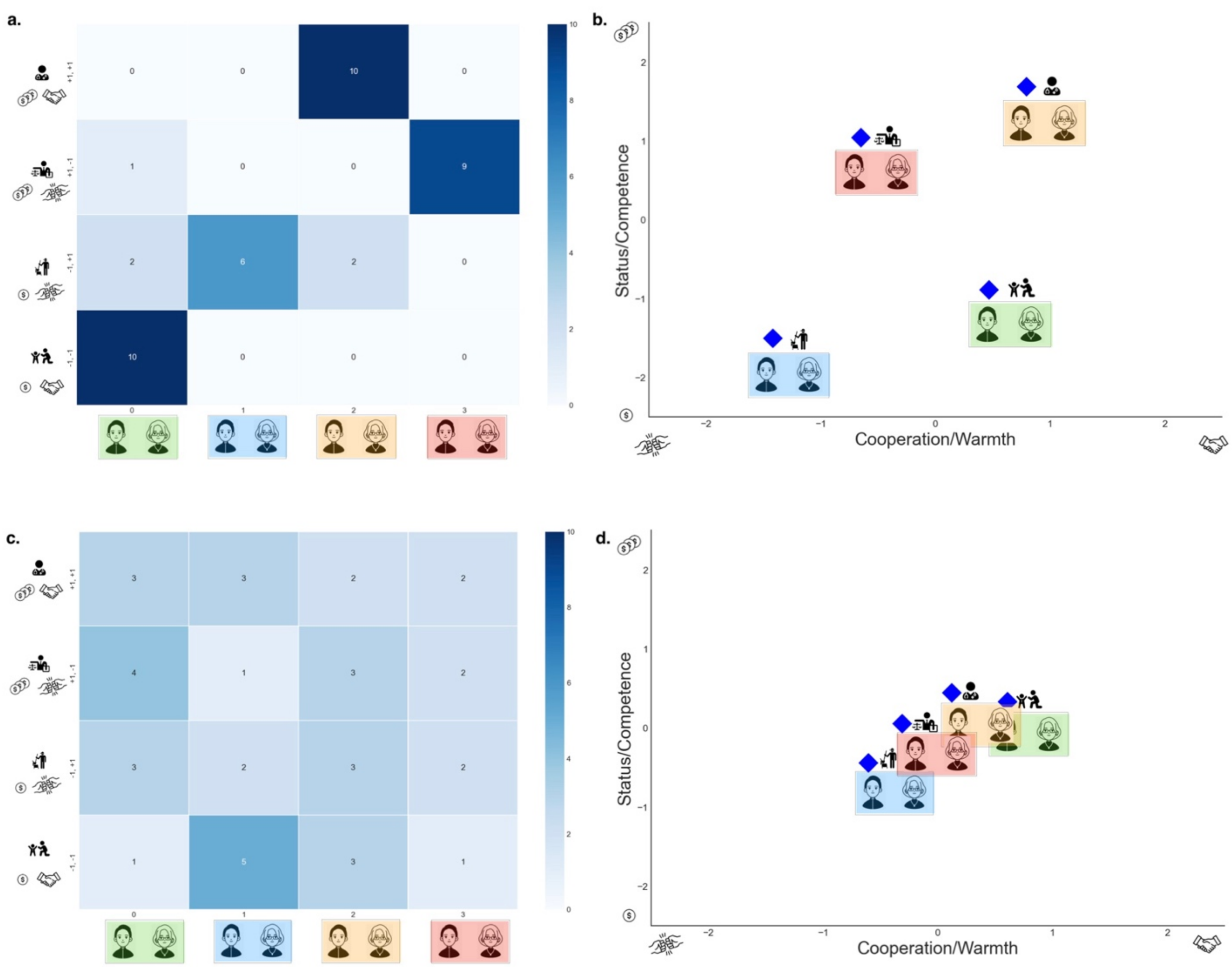

Fig. 2. Two example simulated results from an agent who makes adaptive decisions (in panels a-b) and another agent who makes decisions at random (in panels c-d). The heatmaps on the left panels show how many times a group, on the horizontal axis, is recommended for a job, on the vertical axis. The scatterplots on the right panels show estimated coefficients for the four groups on the two binary features. For aggregate simulation results see *SI* simulation section.

Experiment. We tested the predictions of this model in a large-scale online experiment in which participants ($N = 1310$) 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 (*5, 27*; see *SI* for a preliminary study 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 *SI* for experimental designs).

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 (Fig. 4a. "Default"). In contrast, participants who observe random decisions from the mayor display higher entropy with less stratified and more equally distributed choices (Fig. 4a. "Ideal"). 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], $p < .001$; Fig. 4b. “Default” versus “Ideal”; Fig. 3 for example participants). 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 (*5-6*; see *SI* for more results).

Two results are worth highlighting: First, we see evidence for the emergence of multidimensional stereotypes. As shown in Fig. 3b, 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 (Fig. 3a). Second, we see evidence for generalization. Regardless of the diversity of the jobs, participants clearly find (dis)similarities between jobs. As shown in Fig. 3a, 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. In sum, human behavioral data replicate the model predictions, showing that a stratified society emerges from participants acting adaptively to solve the explore-exploit tradeoff, and that this stratification leads to multidimensional stereotypes with a similar structure to those observed for real social groups.

Fig. 3. Prototypes of stratified vs. diversified hiring choices and dissimilar vs. similar stereotypes. Panels a and b show results from participant #153, who was assigned to the adaptive exploration condition. 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 (a). 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 (b). Panels c and d show results from participant #281, who was assigned to the random exploration condition. This participant observes the mayor selecting randomly (c). As a result, this participant thinks Aimas, Tufas, Wekis, and Rekus are similarly warm and competent (d).

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 explored three interventions to test this prediction: adding an exploration bonus, decreasing the reward probability, and imposing a random holdout. Each intervention addresses the implicit cost of exploration in a different way. First, adding a bonus to untried options directly incentivizes exploration (*34*). 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. Third, randomly holding out some groups to make them unavailable forces exploration, making the cost of exploration irrelevant. 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 *SI* for modeling results). 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 *SI* for experimental designs).

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 (Fig. 4a. "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

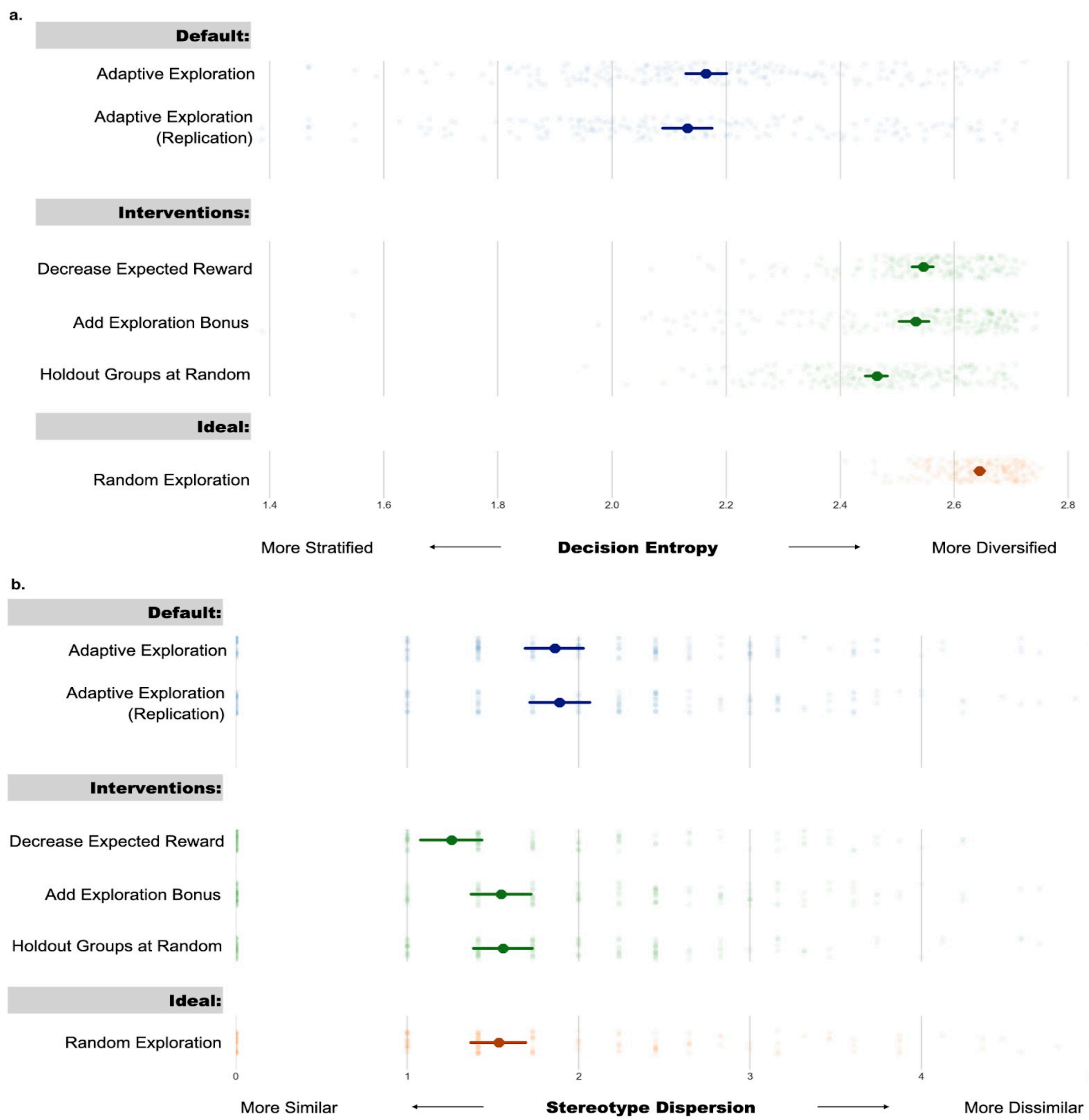

Fig. 4. Average treatment effects in human behavioral experiments. The vertical axis represents experimental conditions: The default panel with blue bars shows the adaptive exploration condition where participants make their own hiring decisions in the main study and replication in the mechanism study. 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. (a) shows more stratified choices to more diversified choices in the order of the default exploration, the interventions, and the random ideal condition. (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.

(*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; Fig. 4b. "Interventions" and "Default replication") conditions. This pattern is robust for future hires and status-cooperation dimensions (see *SI* 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 we have introduced in this paper 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 environmental samples, 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 to the individual because it minimizes exploration costs. Nonetheless, this strategy brings collateral damage to society - because it leaves other groups under-explored 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.

Our proposed mechanism complements but differs from prior theories on the origin of stereotypes, as follows. (a) The motivation to maximize self-interest can be orthogonal to the motivation to maintain group identity or hierarchy (*8-11*). 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 (*34*), 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. (b) A lack of exploration differs from confirmation bias or metacognitive myopia (*12-16*). 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 (*24*). (c) Our proposed mechanism does not depend on asymmetric population sizes when one group is more accessible than other groups (*17-20*). 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. (d) Our proposed mechanism does not endorse stereotype accuracy at all (*21-22*), because we showed 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 (*36*). 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.

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 (e.g., *8*), or the result of power-maintaining or undiversified organizational arrangements (e.g., *37*). 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 under-explore other groups for inexperienced tasks (*38*). A university merely pursuing a higher ranking for research can admit certain kinds of researchers for particular disciplines but under-explore other combinations (*39*). 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 datasets and machine-learning algorithms, propose allocating refugees with similar demographic features to specific locations for similar jobs based on past success (*40*). 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 (*41*). While optimizing customer engagement may be an adaptive strategy for the local algorithm, it simultaneously perpetuates stratification in the global online system.

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 (*42-43*). 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 datasets to examine adaptive exploration in everyday choices (*44-45*). 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 (*46*). 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 here provide a first step in that direction.

Materials and Methods

Experimental details, dataset construction, analysis details, formal modeling, and computational simulations are provided in the Supplementary Information.

References

1. 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.

2. Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *The Journal of Abnormal and Social Psychology*, *28*(3), 280.

3. 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*, *117*(23), 12741-12749.

4. Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623-1626.

5. 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.

6. 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.

7. Bai, X., Fiske, S. T., & Griffiths, T. L. (2022). Globally inaccurate stereotypes can result from locally adaptive exploration. *Psychological Science*, *33*(5), 671-684.

8. Altemeyer, B. (1983). *Right-wing authoritarianism*. Univ. of Manitoba Press.

9. Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate?. *Journal of social issues*, *55*(3), 429-444.

10. 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.

11. 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.

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

13. 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.

14. 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.

15. 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.

16. 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.

17. Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological review*, *107*(4), 659.

18. Denrell, J. (2005). Why most people disapprove of me: experience sampling in impression formation. *Psychological review*, *112*(4), 951.

19. 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.

20. Alves, H., Koch, A., & Unkelbach, C. (2018). A cognitive-ecological explanation of intergroup biases. *Psychological Science*, *29*(7), 1126-1133.

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

22. Jussim, L. (2017). Précis of Social Perception and Social Reality: Why accuracy dominates bias and self-fulfilling prophecy. *Behavioral and Brain Sciences*, *40*, e1.

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

24. 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.

25. 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.

26. Jones, J. M., Dovidio, J. F., & Vietze, D. L. (2013). *The psychology of diversity: Beyond prejudice and racism*. John Wiley & Sons.

27. 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.

28. 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*, *111*(supplement_4), 13593-13597.

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

30. 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.

31. Agrawal, S., & Goyal, N. (2012, June). Analysis of thompson sampling for the multi-armed bandit problem. In *Conference on learning theory* (pp. 39-1). JMLR Workshop and Conference Proceedings.

32. Chapelle, O., & Li, L. (2011). An empirical evaluation of thompson sampling. *Advances in neural information processing systems*, *24*.

33. 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.

34. Bellemare, M., Srinivasan, S., Ostrovski, G., Schaul, T., Saxton, D., & Munos, R. (2016). Unifying count-based exploration and intrinsic motivation. *Advances in neural information processing systems*, *29*.

35. 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.

36. Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237*(4820), 1317–1323.

37. Banaji, M. R., Fiske, S. T., & Massey, D. S. (2021). Systemic racism: individuals and interactions, institutions and society. *Cognitive research: principles and implications*, *6*, 1-21.

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

39. 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.

40. Bansak, K., Ferwerda, J., Hainmueller, J., Dillon, A., Hangartner, D., Lawrence, D., & Weinstein, J. (2018). Improving refugee integration through data-driven algorithmic assignment. *Science*, *359*(6373), 325-329.

41. Dean, S., Rich, S., & Recht, B. (2020, January). Recommendations and user agency: the reachability of collaboratively-filtered information. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 436-445).

42. 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.

43. 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.

44. Card, D., Chang, S., Becker, C., Mendelsohn, J., Voigt, R., Boustan, L., ... & 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*, *119*(31), e2120510119.

45. Charlesworth, T. E., 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*, *119*(28), e2121798119.

46. Almaatouq, A., Griffiths, T. L., Suchow, J. W., Whiting, M. E., Evans, J., & Watts, D. J. (2022). Beyond Playing 20 Questions with Nature: Integrative Experiment Design in the Social and Behavioral Sciences. *Behavioral and Brain Sciences*, 1-55.