Check for updates

Research Article

Psychological Science
1–16
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09567976231180587
www.psychologicalscience.org/PS
Sage

Social Concepts Simplify Complex Reinforcement Learning

Leor M. Hackel[1] and David A. Kalkstein[2]
[1]Department of Psychology, University of Southern California, and [2]Department of Psychology, Stanford University

Abstract
Humans often generalize rewarding experiences across abstract social roles. Theories of reward learning suggest that people generalize through model-based learning, but such learning is cognitively costly. Why do people seem to generalize across social roles with ease? Humans are social experts who easily recognize social roles that reflect familiar semantic concepts (e.g., "helper" or "teacher"). People may associate these roles with model-free reward (e.g., learning that helpers are rewarding), allowing them to generalize easily (e.g., interacting with novel individuals identified as helpers). In four online experiments with U.S. adults ($N = 577$), we found evidence that social concepts ease complex learning (people generalize more and at faster speed) and that people attach reward directly to abstract roles (they generalize even when roles are unrelated to task structure). These results demonstrate how familiar concepts allow complex behavior to emerge from simple strategies, highlighting social interaction as a prototype for studying cognitive ease in the face of environmental complexity.

Keywords
concepts, generalization, reinforcement learning, relational reasoning, rewards, social cognition, open data, open materials, preregistered

Received 9/6/22; Revision accepted 5/9/23

Every social interaction is unique—featuring different people, times, or places—but many interactions have the same general structure. Accordingly, people often generalize rewarding experiences with a particular individual ("Dr. Smith") to a more general social role ("professor"). By doing so, they can learn about rewards available from any individual in that role, such as the value of visiting a professor during office hours before an exam—regardless of the specific professor. These situations exemplify learning across complex *relational structure*, or abstract relationships that hold true across different settings. Despite this complexity, people often seem to interact easily with new individuals in familiar roles. How are people so efficient in using relational structure to generalize reward across social roles?

Past research suggests that generalizing reward across abstract relationships should be effortful. According to this view, people navigate relational structure through model-based reinforcement learning—a cognitively costly strategy in which one uses an internal model of the environment to pursue goals (Daw et al., 2011; Doll et al., 2012, 2015; Kool et al., 2017; Otto et al., 2013). Specifically, people can consult a *cognitive map* that describes abstract structure—how different entities relate to one another or how actions lead to outcomes—allowing people to prospect into the future, infer new paths to reward, and generalize across actions that lead to equivalent outcomes (Behrens et al., 2018; Boorman et al., 2021; Brown et al., 2016; Doll et al., 2015; Karagoz et al., 2022; Park et al., 2021; Wang et al., 2020). For instance, if a New Yorker takes the subway to a friend's neighborhood and enjoys a visit, they can infer that taking a bus to the same neighborhood will lead to the same friend. In contrast, a model-free learner merely repeats actions rewarded in the past, regardless of the

Corresponding Author:
Leor M. Hackel, Department of Psychology, University of Southern California
Email: lhackel@usc.edu

path that led to reward. Compared with model-free learning, model-based learning is cognitively demanding: People are slower to use model-based learning, they need to expend more cognitive resources to do so, and they tend to avoid it unless they are sufficiently motivated (Kool et al., 2017; Otto et al., 2013).

However, people might learn and generalize more easily in social settings. Studies of model-based generalization typically use novel tasks featuring novel relationships, which might particularly strain people's cognitive resources with mental labor (Kool et al., 2018; Otto et al., 2013). Although these studies capture the important task of learning in novel environments, they may leave out more efficient strategies available to people in familiar environments such as social interactions.

Here, we propose such a strategy: People can use familiar semantic concepts to generalize reward across abstract social roles in a model-free manner. People use relational categories (e.g., "helping," "teaching") to identify individuals in familiar relational roles ("helper" and "recipient of help," "teacher" and "student"; Gentner et al., 2011; Gentner & Kurtz, 2005; Goldwater et al., 2011). People may associate these roles directly with reward (e.g., learning that helpers are rewarding), allowing them to generalize easily (e.g., choosing to interact with other individuals identified as helpers). Rather than requiring a learner to plan paths to reward with a cognitive map, this strategy would be more akin to stimulus–response learning, with an abstract role serving as the stimulus. Past work suggests that people learn about abstractions through reinforcement (Diuk et al., 2013; Eckstein & Collins, 2020; Precup & Sutton, 1997), even in the absence of prospection or inference; for instance, people habitually select abstract goals on the basis of reward feedback (Cushman & Morris, 2015). People may similarly learn to choose abstract relational roles directly through reward feedback; although people do often use a model of the world to generalize across social roles (e.g., assuming two financial advisors share expertise), people may also attach reward to roles in a model-free manner. People can thus simplify complex learning problems by using familiar concepts to think about complicated structure.

This strategy may be particularly relevant, though not unique, to social interaction. Humans are social experts with a wealth of semantic concepts that describe social relational roles (e.g., "helper," "adversary," "defendant"; Atzil et al., 2018; Kalkstein, Hackel, & Trope, 2020; Spunt & Adolphs, 2015). In turn, expertise and semantic concepts promote relational reasoning (Gentner, 2016; Gentner et al., 2011; Goldwater et al., 2021; Loewenstein & Gentner, 2005). People may therefore be especially adept at relational reasoning in social domains. Consistent with this proposal, research has found that people recognize and reason about relational structure more easily in social (as opposed to nonsocial) settings (Cosmides, 1989; Kalkstein et al., 2016; Kalkstein, Hackel, & Trope, 2020; Mason et al., 2010). Moreover, although no physical pattern in the world defines social relations (e.g., helping), people are so practiced at recognizing social relations that doing so has hallmarks of automatic perception rather than deliberate reasoning (Hafri & Firestone, 2021). Altogether, although people can use familiar concepts in nonsocial settings, there are few (if any) other domains of familiarity and expertise so universal. Social concepts thus offer a relevant and important test case.

We therefore asked whether people can generalize reinforcement learning across social structures with ease by learning the reward value of abstract roles described by familiar social concepts. We administered sequential (two-step) reinforcement-learning tasks featuring two social scenes with distinct characters in each; characters in different scenes led to the same end states, allowing participants to generalize reward feedback across scenes on the basis of task structure. We manipulated whether characters also embodied roles within the familiar category of "helping." We hypothesized that social concepts would ease complex learning: Participants would

Statement of Relevance

People often need to apply their past social experiences to make choices in new interactions. One way people do this is by generalizing rewarding experiences with individuals ("Lisa") to abstract social roles ("mentors"). Yet past research suggests that this should require slow mental effort as people consult a mental map to identify different individuals who lead to the same outcomes. In this research, we explored whether people also rely on a simpler but less precise strategy to generalize across familiar social roles. People quickly recognize roles described by familiar concepts such as "helper" and may associate these concepts directly with reward, allowing easy generalization to other instances of a concept. Accordingly, we found that learning involving familiar social roles led to quicker and more extensive generalization, including overgeneralization that would not emerge through a mental map. Although effortful reasoning lets people navigate novel abstract roles, our findings suggest that conceptual knowledge simplifies how people navigate familiar roles. This work illuminates how humans make complex social choices with ease.

generalize more and at faster speed when characters embodied social roles. In two additional experiments, we tested whether this pattern emerges because people attach reward directly to abstract roles in a model-free manner instead of planning with a cognitive map. To do so, we tested whether participants would generalize reward across social roles even when doing so was not warranted by task structure—a pattern that would not emerge from model-based inference. These experiments thus tested a key role for social concepts in abstract reinforcement learning.

Open Practices Statement

Sample sizes, measures, exclusion criteria, and analysis plans for Experiments 1b, 2a, and 2b were preregistered. Additional exploratory analyses or deviations are described as such below. Preregistration documents are available at https://aspredicted.org/SPB_PL9 (Experiment 1b), https://aspredicted.org/Q7H_G8Y (Experiment 2a), and https://aspredicted.org/269_BW5 (Experiment 2b). Deidentified data and analysis code have been made publicly available on OSF and can be accessed at https://osf.io/ncdt2.

Experiments 1a and 1b: Social Roles Ease Generalization

Method

Overview. We examined reward generalization in the presence or absence of familiar social roles that were incidental to task structure. Participants completed a sequential (two-step) decision task adapted to depict social scenes (Doll et al., 2015; Hackel et al., 2019; Kool et al., 2017). Participants were told that they would learn about four characters who owned stock in one of two companies—the Brown Parrot Corporation or the Purple Porcupine Company (Fig. 1a). In the first stage of each round, participants saw a pair of characters and chose one to approach. In the second stage, participants saw the chosen character's stock and pressed a button to receive a dividend as a gift. Each character owned one stock throughout the task (i.e., transitions from first-stage choices to second-stage states were deterministic). The task used the logic of generalization to dissociate model-free and model-based learning (rather than using probabilistic transitions as in an alternate variant of the two-step task developed by Daw et al., 2011).

To allow generalization on the basis of task structure, we ensured that characters always appeared in consistent pairings and that each stock was owned by one character in each scene (Figs. 1b and 1c). A learner using model-based prospection could therefore generalize across scenes by considering the stock each character owns: After choosing a character who owns the Brown Parrot stock (e.g., the man in the green shirt) and receiving a large reward, this learner could choose the character in the other scene who owns the same stock (the woman in the blue shirt). In contrast, a model-free learner would fail to generalize across scenes because that learner would have no model indicating that different characters lead to the same stock (Doll et al., 2015; Kool et al., 2017). Instead, a model-free learner would simply attach reward to a particular character ("man in green shirt") and choose that character again. Model-based choices in this task relate to neural markers of prospection (Doll et al., 2015), depend on a cognitive map (Karagoz et al., 2022), and increase when people face higher stakes (Kool et al., 2017).

Critically, each participant was randomly assigned to see one of two versions of the scenes. In the congruent-social-roles condition (present in both Experiments 1a and 1b), each scene displayed one person helping another. The scenes displayed different concrete behaviors, but characters could be aligned across scenes as "helpers" or "recipients of help." Social roles were congruent with task structure: Both helpers owned one stock and both recipients of help owned the other stock. Participants could therefore make choices by tracking which role was rewarding, using roles as a cue to reward. In the no-social-roles condition (Experiment 1a), the scenes showed the same characters standing still with no interaction, and in the incongruent-roles condition (Experiment 1b), social roles were incongruent with task structure. In these conditions, participants therefore had to track which stocks were rewarding and align scenes through model-based prospection (i.e., considering which characters led to rewarding stocks).

Both social roles and stock ownership could thus allow structural alignment (i.e., mentally aligning characters across scenes) and analogical transfer (i.e., transferring reward from a character to their counterpart in the other scene). However, social roles involved familiar concepts, whereas stock ownership involved a novel cognitive map. We hypothesized that social roles would allow greater generalization with lower cognitive cost. Alternatively, traditional approaches suggest that generalizing reward on the basis of abstract relationships is effortful (Kool et al., 2017; Otto et al., 2013). Social roles are abstract, with no one percept defining them. If people use only slow and effortful model-based learning to generalize across abstract relationships, then adding an abstract social role should require additional (rather than less) processing. In this case, familiar concepts would not make learning easier; participants either would show no differences in learning or would be slowed down by considering another abstract role.

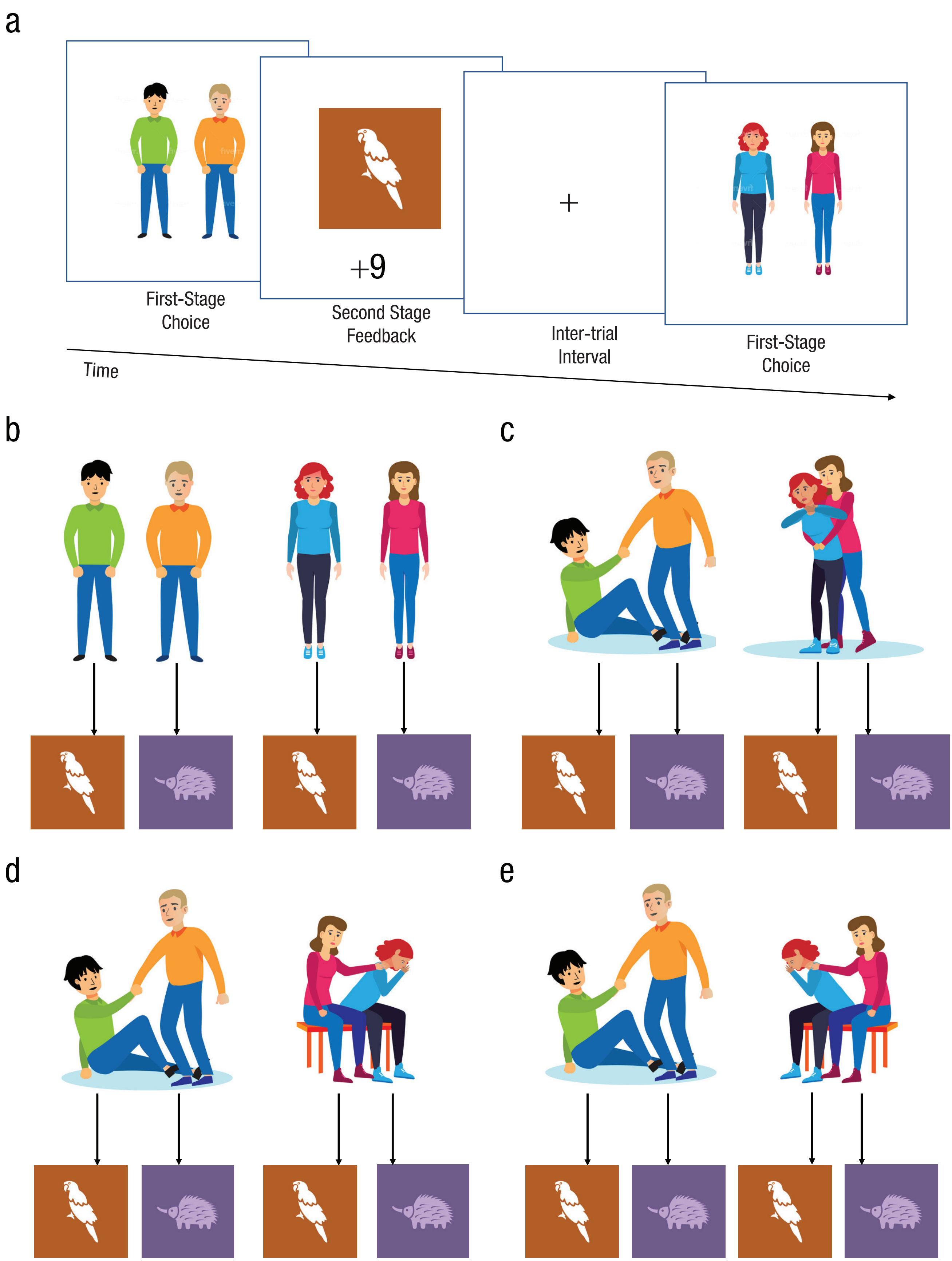
a
+9
First-Stage Choice
Second Stage Feedback
Inter-trial Interval
First-Stage Choice
Time
b
c
d
e

Fig. 1. Stimuli and trial sequence in Experiments 1a and 1b. (a) Participants chose between two characters in a first-stage scene. Choices led to logos representing stocks in the second stage via deterministic transitions; after pressing a button, participants received a reward. On subsequent trials, participants could generalize by choosing the character in the other scene who owned the same stock. In Experiment 1a, each participant was randomly assigned to either (b) a no-social-roles condition, in which characters did not embody relational social roles, or (c) a congruent-social-roles condition, in which characters embodied familiar roles of "helper" (helping up someone who fell, performing the Heimlich maneuver) or "recipient of help." Characters with the same social role owned the same stock. In Experiment 1b, all participants saw characters enacting helping (helping up someone who fell, comforting someone in distress). Each participant was randomly assigned to either (d) an incongruent-social-roles condition, in which characters with the same social role led to different stocks, or (e) a congruent-social-roles condition, in which characters with the same social role owned the same stocks.

Participants. Participants were recruited on the Cloud Research platform and participated in exchange for payment in both Experiment 1a (N = 175; 68 women, 105 men, two preferred not to disclose gender; age: M = 39.21 years, SD = 11.90) and Experiment 1b (N = 277; 122 women, 153 men, two nonbinary; age: M = 40.09 years, SD = 11.50). In Experiment 1a, sample size was chosen to provide 90% power to detect a moderate effect size (d = 0.5). In Experiment 1b, sample size was chosen on the basis of the effect size observed in Experiment 1a. In both experiments, additional participants were recruited to account for exclusions. In Experiment 1b, the preregistered sample size was 275 participants; two additional participants completed the study without requesting payment. Results remained the same when we excluded these two participants. Informed consent was obtained from all participants in accordance with review and approval from the University of Southern California Office for Protection of Human Subjects. For additional information, see the Supplemental Material available online.

To ensure that participants were actively engaged in the task, we administered exclusion rules used in prior work, removing data from any participant who did not respond to at least 80% of first-stage and second-stage choices and who had mean reaction times greater or less than 2 standard deviations from the group mean (Gillan et al., 2015; Hackel & Zaki, 2018; Kool et al., 2017). These criteria left 151 participants for analysis in Experiment 1a and 242 participants in Experiment 1b.

Stimuli. Participants saw two social scenes. In Experiment 1a, each scene either showed a pair of individuals standing still or showed one individual helping the other. In Experiment 1b, we ensured that any resulting effects of condition were not due to other differences in the stimuli; for instance, scenes featuring social interaction might be more engaging than scenes without interaction. To rule out this possibility, we created scenes that always featured a helper and a recipient of help (Figs. 1e and 1f). In the congruent-roles condition, task structure again aligned with social roles; the two helpers owned the same stock, and the two recipients of help owned the same stock. In the incongruent-roles condition, task structure was misaligned with social roles: The helper from each scene owned the same stock as the recipient of help from the other scene. In Experiment 1b, we also replaced one helping scene with a new scene featuring emotional help, thus testing a broader range of behaviors within the abstract category "helping."

Procedure. Participants completed 152 rounds of a sequential decision-making task. On each trial, participants had 1.5 s to choose between two characters in a first-stage scene; participants were told that by approaching these individuals, they would be able to receive a dividend from that person's stock as a gift. Trials began in one of the two social scenes, evenly and randomly distributed throughout the task. On every trial, each character had an equal probability of appearing on the right or left side of the screen.

After choosing a character, participants then observed a transition to a second stage, in which the social scene faded from the screen and a stock logo appeared. The transition lasted from the participant's response until the end of the response window; this meant that each trial lasted for the same amount of time, and participants could not complete the task sooner by responding more quickly. The link between characters and stocks (i.e., which pair of characters led to which logo) was randomized across participants.

Stocks in the second stage were represented by colorful animal stimuli, which participants were told represented logos for different companies. Participants had 1.5 s to press a button in the second stage to reveal a reward. After the button press, a frame appeared around the stock logo until the response window had passed (again ensuring equal trial timing for all participants), at which point the reward feedback was displayed for 1.5 s. Rewards in each trial varied from 0 to 9 points, with Gaussian drift (SD = 2; Kool et al., 2017). One stock was initialized to a random value between 0 and 4 points, and the other was initialized to a random value between 5 and 9 points, with stocks randomly assigned to one of the two ranges. Dividends consisted of points, which served as raffle tickets for a $5 bonus.

Before beginning the task, participants first read extensive instructions, which contained practice trials for each stage of the task (see the Supplemental

Material). After the task, participants were queried for their explicit knowledge about task structure (see the Supplemental Material).

Analytic procedures. To examine the extent of generalization in each condition, we fitted participant choices to a computational model that allows for a hybrid of model-based and model-free learning (Doll et al., 2015; Kool et al., 2017). The model-based component uses the transition structure of the task to identify characters who will lead to desirable stocks and can therefore generalize across scenes. The model-free component chooses characters that previously led to reward without considering the stock each character owns. A weighting parameter w reflects the balance between model-free (w = 0) and model-based (w = 1) control. Although this model formally characterizes generalization through prospection, prospection and role-based learning make identical predictions in this task (i.e., generalization across scenes). We therefore used this model to characterize the degree of generalization in both conditions (using rank-sum tests because of nonnormality of the distribution), allowing us to compare the two conditions with each other and with past work on equal terms. (See the Supplemental Material for further details regarding computational modeling.) We interpreted the w parameter as a measure of generalization across structure (rather than prospection per se).

To characterize the cognitive cost of learning, we examined reaction times during first-stage choices (excluding reaction times faster than 200 ms). Reaction times were fitted to mixed-effects linear regression models with the following predictors: condition (1 = congruent social roles, −1 = no/incongruent social roles), reward feedback (standardized within participants to z scores), and whether each trial began in the same scene or different scene relative to the trial before (1 = same starting scene, −1 = different starting scene), in order to model the effect of task variables that differed across trials.

Results

We found that social roles allowed greater generalization with lower cognitive cost. Participants in Experiment 1a generalized more in the congruent-roles condition (median w = .66) than in the no-roles condition (median w = .38), z = −2.84 p = .004 (rank-sum test; d = 0.50; Fig. 2a). Yet despite generalizing more, participants responded more quickly in the congruent-roles condition (M = 590.68 ms, SD = 106.88) than in the no-roles condition (M = 665.82 ms, SD = 136.93), b = −39.56, SE = 9.41, t(148) = −4.20, p < .001 (d = 0.61; Fig. 2b). Participants who could align task structure with social roles thus generalized to a greater extent while responding more quickly. This finding suggests that these participants did not generalize through model-based prospection, which typically produces slower reaction times relative to model-free control (Otto et al., 2013), and instead used the less costly strategy of choosing rewarding roles.

Participants in the congruent-roles condition also displayed lower switch costs in reaction time. When a model-free learner sees different scenes across two trials, they must adapt to the new scene and may respond more slowly than when seeing the same scene twice. In contrast, a learner who focuses perfectly on abstract structure should have no switch costs; this learner would view the two scenes as identical because both scenes have the same abstract structure (Doll et al., 2015). When scenes switched across trials, participants in the congruent-roles condition slowed less than those in the no-roles condition, as revealed by an interaction of condition and starting scene, b = 5.15, SE = 2.47, t(148.39) = 2.08, p = .04. This result suggests that participants who saw social roles were more robustly able to mentally align the scenes and respond on the basis of abstract structure.

The results of Experiment 1b ruled out the possibility that this pattern was due to other differences in the stimuli across conditions, such as engagement due to the presence or absence of social interaction. All participants in Experiment 1b viewed the same stimuli, but social roles were either congruent or incongruent with task roles. When social roles did not align with task structure, participants would need to use model-based prospection to succeed in the task (e.g., "choose the character who leads to the brown stock"). In contrast, when social roles aligned with task structure, participants could switch to an easier role-based strategy (e.g., "choose helpers"). Consistent with this view, results showed that the weighting parameter w was greater in the congruent-roles condition (Mdn = .67), than in the incongruent-roles condition (Mdn = .35), z = −5.48, p < .001 (rank-sum test; d = 0.76). Participants in the congruent-roles condition again responded more quickly overall, b = −28.50, SE = 6.68, t(239.67) = −4.26, p < .001 (d = 0.51), and showed lower switch costs in reaction time, b = 9.67, SE = 1.98, t(235.80) = 4.90, p < .001. These findings rule out engagement as an alternate explanation, instead supporting the proposal that participants generalized by focusing on social roles.

Experiments 2a and 2b: Overgeneralization Across Familiar Social Roles

Method

Overview. In Experiments 1a and 1b, participants generalized more often and more easily across scenes when

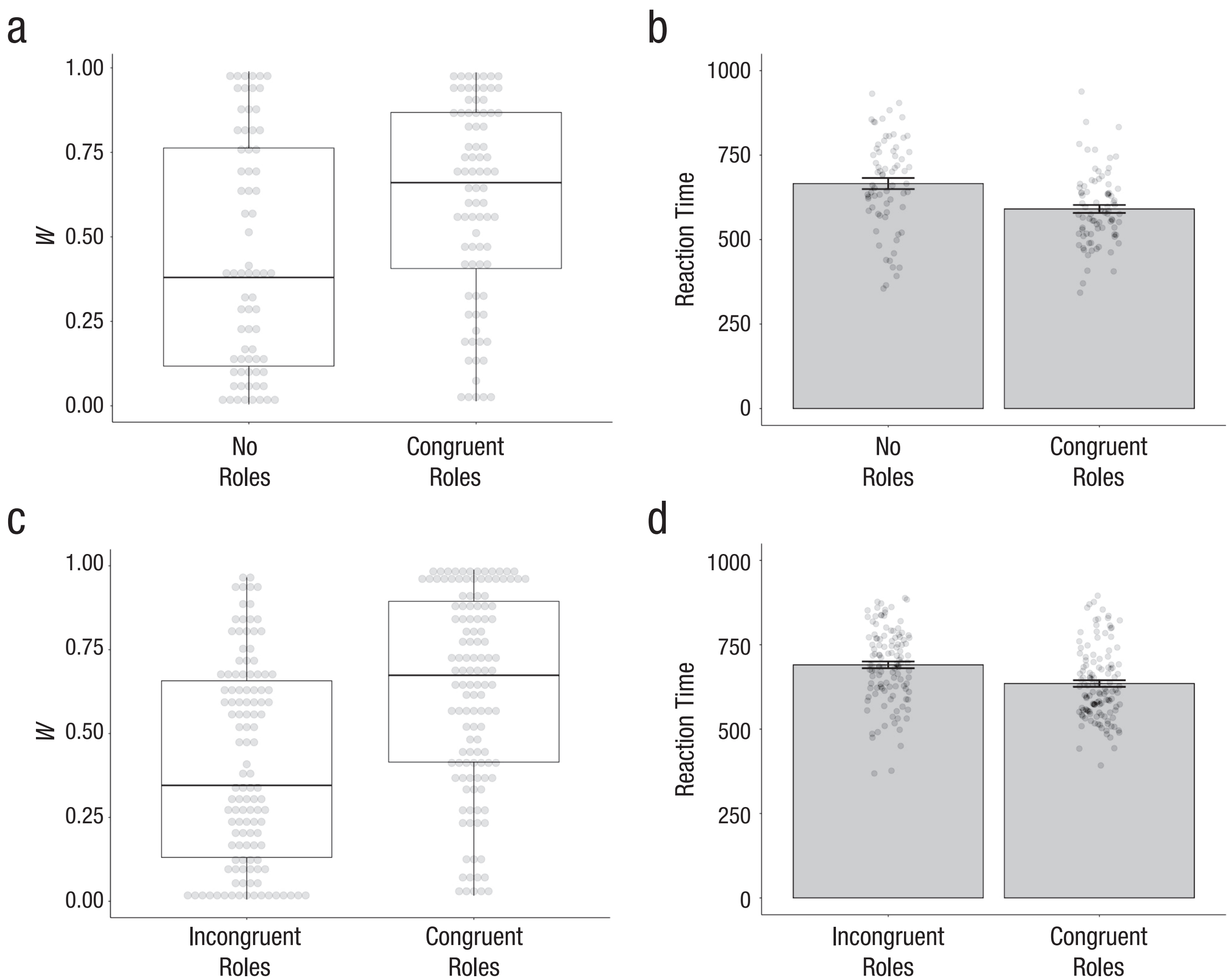

Fig. 2. Experiments 1a and 1b results across choice and reaction time. (a) Degree of generalization based on task structure is indicated by a weighting parameter (w), which reflects a learning focus on stimuli ($w = 0$) versus abstract structure ($w = 1$). This parameter was higher in Experiment 1a—indicating greater generalization—among participants who viewed social roles as opposed to those who did not. Boxes indicate interquartile ranges, whiskers indicate minimum and maximum points, central lines indicate medians, and circles show individual data points. (b) Reaction times during first-stage choice in Experiment 1a were faster among participants who viewed social roles as opposed to those who did not. Error bars indicate standard errors of the mean. Dots indicate raw data points (with jitter), with darker shades representing greater density of data points. (c) Degree of generalization in Experiment 1b. Participants generalized more across structure when social roles were congruent rather than incongruent with task roles. Boxes indicate interquartile ranges, whiskers indicate minimum and maximum points, central lines indicate medians, and circles show individual data points. (d) Reaction times during first-stage choices in Experiment 1b. Participants responded more quickly when social roles were congruent rather than incongruent with task roles. Error bars indicate standard errors of the mean. Dots indicate raw data points (with jitter), with darker shades representing greater density of data points.

they could focus on familiar roles instead of novel transitions—a quantitative shift in generalization. This pattern is consistent with the proposal that participants used different strategies in each condition—learning about the reward value of roles versus prospecting about stocks. However, Experiments 1a and 1b did not qualitatively dissociate these strategies, given that both strategies predict generalization. These studies leave ambiguous whether familiar social roles truly led participants to adopt a different learning strategy or whether social roles simply facilitated model-based prospection. We therefore conducted Experiments 2a and 2b to test our proposed mechanism—namely, that people associated social roles directly with reward.

To do so, we tested qualitatively different predictions of model-free learning, model-based prospection, or role-based learning. Participants again saw two scenes in which social roles aligned with task structure: Two helpers owned the same stock, and two recipients of help owned the same stock. However, we introduced a third scene featuring a helper and recipient of help in which each character owned a unique stock; the task structure in this scene was thus unrelated to the task structure of the first two scenes (Fig. 3a). This design allowed qualitatively different predictions for model-free, model-based, or role-based learning, rather than the quantitative shifts observed in Experiments 1a and 1b. A model-free learner again would not generalize reward across scenes at all, because they associate reward with individual characters (e.g., "choose the man in the green shirt"). Even though transitions are deterministic, this learner does not know that "woman in the blue shirt" owns the same stock, nor does it know that both characters have the same social role. As a result, receiving a reward would lead only a model-free learner to choose the same exact character in the exact same scene again (but not characters in any other scenes).

A model-based learner using prospection would generalize reward only across the two scenes that share task structure because they make choices on the basis of transitions to stocks (e.g., "choose someone who owns the brown stock"). Given that transitions are deterministic, they know that if "man in green shirt" led to a large reward, then "woman in blue shirt" will likely lead to a large reward, as she is guaranteed to lead to the same stock. However, a model-based learner using task structure would not care about whether characters are depicted as helpers or recipients of help, because this is irrelevant to the stocks they own. As a result, receiving a reward would lead a model-based learner to choose characters who own the same stock again (but not characters who own different stocks).

Finally, a role-based learner would generalize reward across all three scenes, because they make choices on the basis of social roles (e.g., "choose a helper"). A role-based learner construes characters in terms of their roles, seeing each action as "choose helper" or "choose recipient of help." The role-based learner does not use knowledge about stocks; they simply learn which abstract role leads to reward (choosing helpers or recipients of help). Because each scene has a helper and a recipient of help, the role-based learner generalizes across all three scenes.

Altogether, "the green character" appears only in one scene, "someone who owns the brown stock" appears in two scenes, and "helpers" appear in all three scenes. Depending on which representation a learner uses, they would apply reward feedback to characters in one, two, or three scenes. This experiment thus let us examine how much participants relied on each strategy. We hypothesized that participants would generalize on the basis of social role, even when doing so was not relevant to the task structure. This finding would indicate that people learn to associate social roles directly with reward. Alternatively, traditional approaches suggest that people generalize across abstract relationships using a cognitive map of task structure; if participants use social roles only to facilitate planning within task structure, then they would generalize across social roles only when relevant to task structure.

Participants. Participants were recruited on the Cloud Research platform and participated in exchange for payment in both Experiment 2a (N = 60; 31 women, 29 men; age: M = 37.32 years, SD = 9.68) and Experiment 2b (N = 65; 25 women, 39 men, one preferred not to disclose gender; age: M = 38.97 years, SD = 11.28). In Experiment 2a, sample size was set heuristically given the within-participants design and large number of trials. In Experiment 2b, sample size was set on the basis of a simulation-based power analysis using Experiment 2a data; additional participants were recruited to account for potential exclusions. We applied the same exclusion rules as in Experiments 1a and 1b. In Experiment 2a, the second exclusion rule (regarding reaction times) was left out of the preregistration because of an error. We report analyses keeping both rules for consistency across studies; however, using only the first rule regarding response rates in Experiment 2a did not change the results. These criteria left 57 participants in Experiment 2a and 60 participants in Experiment 2b.

Procedure. The task design was similar to that in Experiments 1a and 1b, with the following changes. First, participants learned about three social scenes featuring helpers and recipients of help. For two scenes, social roles were congruent with task structure: Both helpers owned the same stock, and both recipients of help owned the same stock. For a third scene, however, social roles were unrelated to task structure: Both the helper and recipient of help owned unique stocks (Fig. 3a). This addition allowed us to ask whether participants would generalize reward across social roles even when roles were unrelated to task structure.

In Experiment 2a, all scenes featured characters of the same gender in order to reduce variability associated with social categories, and scenes were randomly assigned to different stock contingencies across participants (e.g., having shared or unique stocks). In Experiment 2b, to provide a more conservative test of role-based learning, we made the scene with unrelated task structure visually distinct from the others. Specifically, this scene featured female characters, whereas the two scenes with shared structure featured male characters (Fig. S1 in the Supplemental Material). This gender difference made it easier for participants to immediately recognize that this scene had a distinct task structure (see the Supplemental Material).

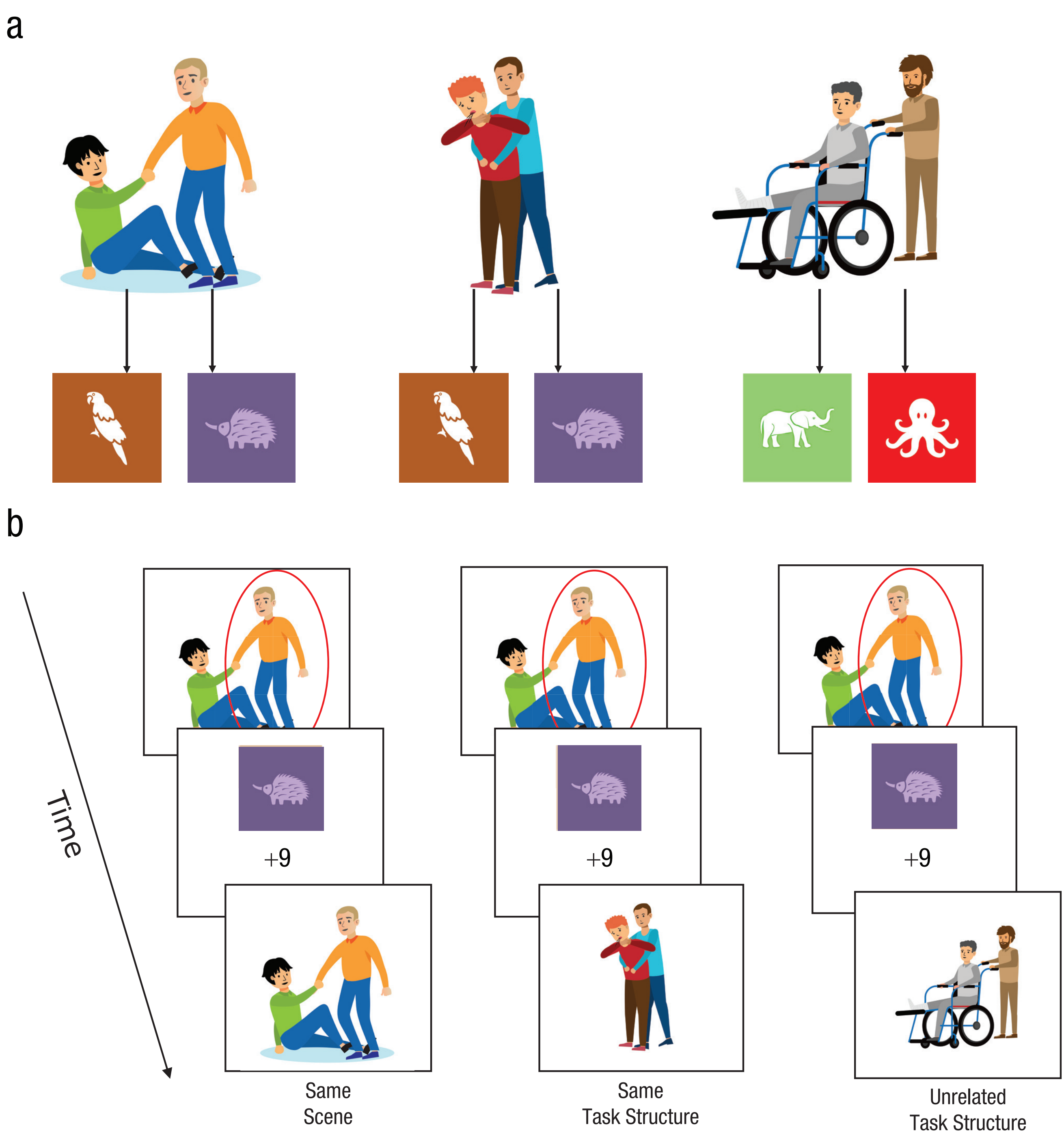

Fig. 3. Stimuli and trial sequence in Experiment 2a. (a) In two scenes, characters with the same social role led to the same stocks. In a third scene, characters led to unique stocks. (b) This design produced three trial types that dissociated learning strategies. Across any two trials in the task, participants could see the same scene twice in a row, different scenes with the same task structure, or different scenes with unrelated task structure. After choosing a helper (designated here by a red circle) and receiving a large reward, a model-free learner would choose a helper on the next trial only when seeing the same scene again. A model-based learner using prospection would also choose a helper when seeing a different scene with the same task structure. A role-based learner would choose a helper again even when seeing a different scene with an unrelated task structure.

Participants completed 180 trials of the learning task. To ensure that participants would see each trial type enough times, we pregenerated 25 pseudorandom stimulus orders in which there were at least 36 trials each for same-scene, same-task-structure, and different-task-structure trial types. Each participant was randomly assigned to one of these 25 stimulus orders. The rewards provided by each stock again fluctuated independently over the course of the task, as in Experiments 1a and 1b, such that the rewards provided by one stock were uninformative about the rewards provided by another stock. After each experiment, participants were asked about the transition structure and reward structure of the task (see the Supplemental Material).

Analytic procedures. Our primary analyses examined qualitative patterns of choice, asking whether large rewards on one trial led participants to choose a character embodying the same social role again on the next trial (Fig. 3b). We asked whether this tendency depended on whether both trials depicted the same scene, different scenes with the same task structure (i.e., the same social roles leading to the same stocks), or different scenes with unrelated task structure (i.e., the same social roles leading to unique stocks). To do so, we fitted first-stage choices to a lagged mixed-effects logistic regression model predicting, on a trial-by-trial basis, whether participants repeated their most recent choice of social roles (1 = stay, 0 = switch). Predictors were the reward earned on the previous trial (standardized within participants to z scores) and whether the previous trial started in the same scene, a different scene with the same task structure, or a different scene with an unrelated task structure. This model approximated the full reinforcement-learning model by asking whether participants stay with their previous choice as a function of feedback and transition structure (Doll et al., 2015; Otto et al., 2013). We fitted the model using different dummy-coding schemes to examine different patterns of interest (see the Supplemental Material for more detail). Models were fitted using the *lme4* package (Bates et al., 2015) in *R* (R Core Team, 2016). Random variances were allowed for the intercept and all slopes that varied within participants (see Table S3 in the Supplemental Material for all coefficients).

Results

Participants used a mix of model-free learning, model-based prospection, and role-based learning (Fig. 4a). Consistent with model-free learning, participants tended to choose the same social role after reward across consecutive trials featuring the same scene (main effect of reward in same-scene trials: $b = 1.25$, $SE = 0.08$, $z = 15.36$, $p < .001$). Consistent with model-based prospection, participants also chose the same social role after reward across trials featuring different scenes with a shared task structure (main effect of reward in same-task-structure trials: $b = 0.82$, $SE = 0.07$, $z = 11.68$, $p < .001$); a slight decline in the effect of reward highlighted distinct contributions of model-free learning and model-based prospection ($b = 0.51$, $SE = 0.09$, $z = 5.50$, $p < .001$). Consistent with role-based learning, participants chose the same social role after reward more often than chance even across trials that featured different scenes with unrelated task structures (main effect of reward in unrelated-task-structure trials: $b = 0.33$, $SE = 0.05$, $z = 6.19$, $p < .001$); a decline in the effect of reward highlighted distinct contributions of model-based prospection and role-based learning ($b = 0.56$, $SE = 0.09$, $z = 6.54$, $p < .001$).

Supporting these analyses, we fitted participant choices to three computational models of learning. The first was the standard hybrid model of model-free and model-based learning described in Experiments 1a and 1b. The second model added a third learning strategy in which learners construe characters as social roles and associate roles directly with reward. Specifically, after choosing a stimulus, model-free learning was generalized to other stimuli that embodied the same social role, scaled by a generalization parameter g between 0 (no generalization) and 1 (full generalization). In a third model, the generalization parameter was also applied to perseveration, or a tendency to stick with the same choice across trials regardless of reward feedback; participants who construe characters as social roles might stick with the same social role across trials (see the Supplemental Material for more detail). The full model including role-based learning and role-based perseveration provided the best fit to behavior, protected exceedance probability = .95 (median $g = .51$). These findings replicated in Experiment 1b, across regression analyses and computational modeling (Fig. 4c; see also Tables S3–S5 in the Supplemental Material).

We next examined reaction times during first-stage choices, hypothesizing that reaction times would vary by trial type. (These analyses were exploratory in Experiment 2a and preregistered in Experiment 2b.) When participants saw the same scene across two trials, they could choose or avoid the same character, allowing fast responses. When participants saw a different scene with the same task structure as in the previous trial, they may have experienced a switch cost, leading to slower responses. Finally, when participants saw a scene with unrelated task structure relative to the previous trial, prospection and role-based learning would conflict, leading to even slower responses. To test these predictions, we fitted reaction times to a mixed-effects linear regression model using the same predictors as in the analysis of choice (again excluding reaction times faster than 200 ms).

Consistent with these hypotheses, reaction times in Experiment 1a were fastest for same-scene trials ($M = 622$ ms, $SD = 102$), slower for same-task-structure trials

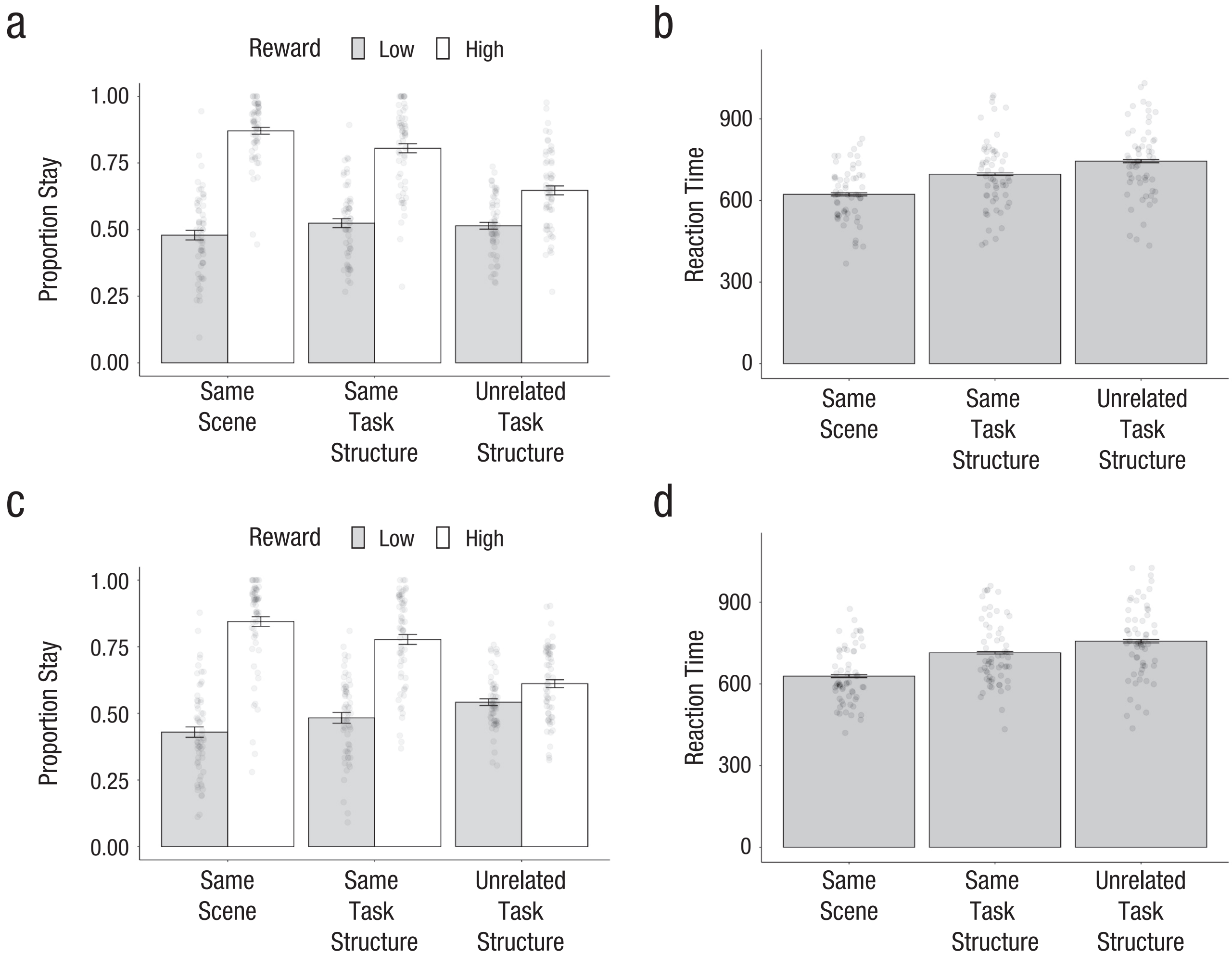

Fig. 4. Experiment 2a and 2b results across choice and reaction time. (a) Qualitative patterns of choices in Experiment 2a revealed contributions of model-free learning, model-based prospection, and role-based learning. After receiving a large reward, participants were most likely to stay with the same social role on the next trial if they saw the same scene again, a little less likely to stay with the same role if they saw a different scene with the same task structure, and less likely—but more likely than chance—to stay with the same role when they saw a scene with an unrelated task structure relative to the previous trial. High and low reward corresponds to reward values below or above the midpoint of the reward distribution. (b) Reaction times in Experiment 2a were fastest when the same scene was viewed across two trials, slower when different scenes featuring the same task structure were viewed across two trials, and slowest when different scenes featuring unrelated task structure were viewed across two trials. Experiment 2b replicated these patterns across (c) choice and (d) reaction time. In all panels, error bars indicate standard errors of the mean with within-participants adjustment (Morey, 2008), and dots indicate raw data points (with jitter), with darker shades representing greater density of data points.

(M = 697 ms, SD = 131), b = 75.28, SE = 7.59, t(59.81) = 9.92, p < .001, and slowest for unrelated-task-structure trials (M = 745 ms, SD = 135), b = 121.10, SE = 9.14, t(59.91) = 13.25, p < .001 (Fig 4b; see also Table S6 in the Supplemental Material). These findings again replicated in Experiment 2b (Fig. 4d; Table S6).

Exploratory analyses further supported the idea that role-based learning promotes cognitive ease. For each participant, we computed the probability that they would select their chosen option on each trial under models that used only model-free, model-based, or role-based learning (but otherwise using each participant's best-fitting parameters). This approach gives a trial-by-trial index reflecting how consistent each choice was with each form of learning (Duncan et al., 2018). We asked how reaction times changed when participants made higher-value (as opposed to lower-value) choices under each value representation. Reaction times were analyzed using mixed-effects regression, with all three value terms serving as predictors in one model (standardized within participants to z scores). For model-based values, the relationship between value and reaction time was small and nonsignificant, $b = -6.84$, $SE = 4.40$, $t(93.07) = -1.56$, $p = .12$; that is, reaction times were not significantly faster when participants made choices with higher versus lower model-based values, consistent with the idea that model-based learning requires slow inference. In contrast, reaction times were significantly faster when choices were more consistent with model-free value, $b = -20.61$, $SE = 3.71$, $t(82.75) = -5.56$, $p < .001$, or with role-based value, $b = -22.53$, $SE = 3.42$, $t(63.80) = -6.59$, $p < .001$. A contrast of coefficients for role-based versus model-based value indicated that these coefficients differed significantly from one another, $\chi^2(1) = 6.57$, $p = .01$, consistent with the idea that distinct cognitive processes supported role-based and model-based choice. (Experiment 2b made model-based learning easier by providing gender as a visual cue to task structure, thus changing the dynamics of model-based learning by reducing its cognitive cost—a change reflected in significantly higher w parameters, as described in the Supplemental Material. Nonetheless, the same qualitative pattern emerged, and analyzing the combined data of Experiments 2a and 2b yielded identical inferences; see the Supplemental Material, including Table S8).

Altogether, we found that participants made choices through a combination of model-free learning (choosing characters that previously led to rewards), model-based prospection (choosing characters that led to rewarding stocks), and role-based learning (choosing social roles that previously led to reward). Participants thus construed characters in terms of abstract social roles and chose accordingly.

General Discussion

Social interactions are complex (FeldmanHall & Nassar, 2021), but humans are adept at navigating them. Here, we identified a reinforcement-learning strategy that allows humans to generalize over abstract social structure with ease: People directly learn the reward value of relational roles described by familiar semantic concepts (e.g., "helpers"), allowing them to generalize easily across different individuals who embody those roles.

This finding identifies a route to reward generalization without cognitively costly inference or planning. When people mentally align two scenes—recognizing how features of one map onto features of a second—they can transfer knowledge across them (Gick & Holyoak, 1983; Holyoak, 2012). Although alignment can stem from representing paths to a goal in a cognitive map (e.g., "Peter and Ana both know a lot about San Francisco and can give helpful advice about visiting"), alignment can also stem from recognizing familiar roles in relational categories (e.g., "Peter and Ana are both tour guides"). Alignment through familiar concepts enabled relatively efficient learning; participants made faster choices when social roles aligned with task structure. At the same time, this strategy led to overgeneralization; participants chose social roles even when roles were unrelated to task structure. Altogether, whereas past work suggests that people use an effortful cognitive map to generalize across abstract relationships (Karagoz et al., 2022), we found that people can use simpler reward association when construing stimuli in terms of familiar concepts.

These findings suggest a continuum of reinforcement learning across levels of abstraction. At one end, people learn to repeat concrete actions ("choose the option on the left"; Shahar et al., 2019) or choices of stimuli ("choose Peter"; Daw et al., 2011; Doll et al., 2015). Choosing a stimulus—as in traditional model-free learning—requires abstracting object identity across time and space, but this abstraction is relatively inflexible because it is tied to one entity. At the other extreme, people learn to repeat actions that lead to an end state or a goal, as in model-based learning rooted in prospection (Doll et al., 2015; Kool et al., 2017). Model-based prospection is a form of abstraction because it creates classes of equivalency based on whether stimuli lead to a given end state (e.g., "choose things-that-lead-to-the-red-room")—a feature that makes it particularly flexible, given that any two stimuli can be linked ad hoc on the basis of environmental structure and goals (Barsalou, 1983; Gilead et al., 2019; Rosch et al., 1976). Here, we identify a strategy wherein people repeat choices of abstract relational roles, akin to model-free learning with an abstract role as the input. This strategy offers more flexibility than traditional model-free learning because it transcends particular stimuli, but it offers less flexibility than model-based prospection because it is tied to preexisting (rather than ad hoc) relationships. Together with prior work demonstrating habitual goal selection (Cushman & Morris, 2015), this work highlights different ways in which abstract concepts support reinforcement learning.

An alternative interpretation of Experiments 2a and 2b is that participants did use model-based inference to generalize across roles, building a model of the task in which stocks owned by characters in the same role were correlated with one another. We believe that this interpretation is less likely for four reasons. First, participants did not *experience* correlated stocks; different stocks drifted independently, meaning that a model-based learner should assign predictive value to stocks rather than to roles. Second, participants did not *expect* correlated stocks; they were carefully instructed that stocks would fluctuate independently and had to correctly answer a quiz question affirming this point; results further remained the same when we restricted analyses to participants who reported after the task that stocks were uncorrelated (and who had perfect knowledge of the transition structure; see the Supplemental Material). Third, in Experiment 2a, participants were relatively slow to make choices based on task structure—consistent with the idea that model-based inference requires slow effort—but faster to make choices based on model-free or role-based values. Fourth, the best-fitting model included perseveration based on roles; participants were likely to stick with the same role across trials regardless of reward outcome. This finding suggests they construed actions in terms of roles ("choose helpers" vs. "choose recipients of helps") rather than using roles solely to reason about which individual characters owned rewarding stocks. (For further discussion of related interpretations, see the Supplemental Material).

At the same time, in Experiments 1a and 1b, it is likely that participants did use their model of task structure to recognize that helpers owned the same stock, after which they learned in terms of the social roles. People may thus use a cognitive map to initially determine which concepts are relevant, after which they can use the concepts during learning without referring back to the cognitive map. Future work can further explore this intersection of learning processes.

Role-based learning may help people navigate social systems, letting people easily learn how to interact with different individuals who occupy identical roles as competitors, peer reviewers, or assistants. In the present work, social roles were not inherently related to task structure, allowing us to experimentally dissociate the influence of each. However, in daily life, it is likely that people use relational categories because these categories usually do relate to reward structure; in a medical emergency, the role of "doctor" can lead to reward, and when a worker needs to borrow a stapler, a "helper" can lead to reward. Reward learning may therefore be intertwined with concept development. As people compare social scenarios and identify common features that lead to reward, they may develop relational categories (Gentner & Kurtz, 2005). These categories can then ease future reward generalization. This process is distinct from stereotypes, which are thought to reflect semantic associations rather than reward associations (Amodio, 2019). To generalize reward feedback based on stereotypes, people might use semantic knowledge in model-based planning (e.g., "I think this group is trustworthy, therefore they are likely to repay a loan"). In contrast, the present work suggests that people directly associate a social role itself with reward. Although we expect these findings to apply to anyone who quickly recognizes a given role, these experiments were conducted with adults in the United States, and future work can test the extent to which these findings generalize to other samples.

Although we examined relational categories, other kinds of categories might also facilitate generalization, such as categories describing social groups (Dunsmoor & Murphy, 2014; Hackel et al., 2022; Kalkstein, Bosch, & Kleiman, 2020; Osherson et al., 1990). And although we examined social structure, people may easily generalize reward in nonsocial settings whenever they have expertise and fluency with structure (Goldwater et al., 2021). For instance, a New Yorker might generalize across subway lines more easily than a tourist who has recently memorized the subway map. Nonetheless, although people can learn through familiar concepts in any domain—whether social or nonsocial—social settings offer a useful prototype. First, social interaction is a domain in which most people are deeply familiar with abstract concepts; humans are social experts who readily recognize relational structure in social settings (Hafri & Firestone, 2021; Kalkstein, Hackel, & Trope, 2020; Mason et al., 2010). Second, social interaction is a crucial domain for survival and well-being, yet it is a domain in which people face particularly complex learning tasks (FeldmanHall & Nassar, 2021). As a result, it is important to understand how people navigate that complexity.

By identifying a role for familiarity and conceptual ease in reinforcement learning, the present findings raise new perspectives on model-based control in humans. Recent work suggested that humans primarily use model-based learning and that model-free learning observed in past experiments reflects confusion about instructions; when participants viewed expanded instructions that clarified each element of a task, they were primarily model based (Feher da Silva & Hare, 2020). However, in Experiments 1a and 1b of the present work, all participants read identical instructions; clarity of instructions therefore cannot account for participants' lack of generalization when roles did not align with task structure. These results raise the possibility that conceptual difficulty, rather than misunderstanding of instructions, hampered learning in past work. Indeed, people learn contingencies more easily when tasks are framed with meaningful and familiar labels as opposed to abstract and arbitrary labels (Camerer, 1981; Murphy & Medin, 1985; Wright & Murphy, 1984). Experiments that explain all elements of a task with an intuitive narrative (Feher da Silva & Hare, 2020) or that give participants extensive practice (Economides et al., 2015) may increase model-based control by increasing the conceptual ease of using the model.

More broadly, the present work suggests an important interface between reinforcement learning and conceptual knowledge. Studies of human reinforcement learning often use novel scenarios and stimuli, illuminating how people learn about unfamiliar structure—a key challenge for humans to initially make sense of the world or to navigate novel settings. However, real-life learning often involves familiar relations that people recognize fairly automatically (Hafri & Firestone, 2021) and familiar concepts that help people make sense of new information (Murphy & Medin, 1985). People may therefore generalize over familiar structures more easily than would be predicted by previous studies of model-based control.

In short, we found that people directly learn the value of abstract roles described by familiar social concepts, allowing them to generalize reinforcement learning across social structure with low cognitive cost. A key source of human social success may lie in the social concepts we develop early in life (Atzil et al., 2018); these concepts help people become social experts who easily perceive, reason about, and make decisions involving abstract social relations (Cosmides, 1989; Hackel et al., 2020; Heider, 1958; Kalkstein, Hackel, & Trope, 2020; Mason et al., 2010; Read, 1987; Todorov & Uleman, 2002; Winter & Uleman, 1984). Social interaction may thus offer a prototype for studying cognitive ease in the face of environmental complexity, highlighting how the initial work of concept development can ease the later work of complex decision-making.

Transparency

Action Editor: Yoel Inbar
Editor: Patricia J. Bauer
Author Contributions

Leor M. Hackel: Conceptualization; Formal analysis; Investigation; Methodology; Writing – original draft; Writing – review & editing.

David A. Kalkstein: Conceptualization; Methodology; Writing – original draft; Writing – review & editing.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Acknowledgments

We thank Payam Piray and members of the University of Southern California Social Learning & Choice Lab for helpful comments.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/09567976231180587

References

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

Atzil, S., Gao, W., Fradkin, I., & Barrett, L. F. (2018). Growing a social brain. *Nature Human Behaviour*, *2*(9), 624–636. https://doi.org/10.1038/s41562-018-0384-6

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, *11*(3), 211–227.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron*, *100*(2), 490–509.

Boorman, E. D., Sweigart, S. C., & Park, S. A. (2021). Cognitive maps and novel inferences: A flexibility hierarchy. *Current Opinion in Behavioral Sciences*, *38*, 141–149.

Brown, T. I., Carr, V. A., LaRocque, K. F., Favila, S. E., Gordon, A. M., Bowles, B., Bailenson, J. N., & Wagner, A. D. (2016). Prospective representation of navigational goals in the human hippocampus. *Science*, *352*(6291), 1323–1326.

Camerer, C. F. (1981). *The validity and utility of expert judgment*. The University of Chicago.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, *31*(3), 187–276.

Cushman, F., & Morris, A. (2015). Habitual control of goal selection in humans. *Proceedings of the National Academy of Sciences, USA*, *112*(45), 13817–13822.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215.

Diuk, C., Tsai, K., Wallis, J., Botvinick, M., & Niv, Y. (2013). Hierarchical learning induces two simultaneous, but separable, prediction errors in human basal ganglia. *Journal of Neuroscience*, *33*(13), 5797–5805.

Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D., & Daw, N. D. (2015). Model-based choices involve prospective neural activity. *Nature Neuroscience*, *18*(5), 767–772.

Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology*, *22*(6), 1075–1081.

Duncan, K., Doll, B. B., Daw, N. D., & Shohamy, D. (2018). More than the sum of its parts: A role for the hippocampus in configural reinforcement learning. *Neuron*, *98*(3), 645–657.

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

Eckstein, M. K., & Collins, A. G. (2020). Computational evidence for hierarchically structured reinforcement learning in humans. *Proceedings of the National Academy of Sciences, USA*, *117*(47), 29381–29389.

Economides, M., Kurth-Nelson, Z., Lübbert, A., Guitart-Masip, M., & Dolan, R. J. (2015). Model-based reasoning in humans becomes automatic with training. *PLOS Computational Biology*, *11*(9), Article e1004463. https://doi.org/10.1371/journal.pcbi.1004463

Feher da Silva, C., & Hare, T. A. (2020). Humans primarily use model-based inference in the two-stage task. *Nature Human Behaviour*, *4*(10), 1053–1066.

FeldmanHall, O., & Nassar, M. R. (2021). The computational challenge of social learning. *Trends in Cognitive Sciences*, *25*(12), 1045–1057.

Gentner, D. (2016). Language as cognitive tool kit: How language supports relational thought. *American Psychologist*, *71*(8), 650–657. https://doi.org/10.1037/amp0000082

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure mapping and relational language support children's learning of relational categories: Structure mapping and relational language. *Child Development*, *82*(4), 1173–1188.

Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin* (pp. 151–175). American Psychological Association. https://doi.org/10.1037/11156-009

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38.

Gilead, M., Trope, Y., & Liberman, N. (2019). Above and beyond the concrete: The diverse representational substrates of the predictive brain. *Behavioral and Brain Sciences*, *43*, Article e121. https://doi.org/10.1017/S0140525X19002000

Gillan, C. M., Otto, A. R., Phelps, E. A., & Daw, N. D. (2015). Model-based learning protects against forming habits. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(3), 523–536.

Goldwater, M. B., Gentner, D., LaDue, N. D., & Libarkin, J. C. (2021). Analogy generation in science experts and novices. *Cognitive Science*, *45*(9), Article e13036. https://doi.org/10.1111/cogs.13036

Goldwater, M. B., Markman, A. B., & Stilwell, C. H. (2011). The empirical case for role-governed categories. *Cognition*, *118*(3), 359–376.

Hackel, L. M., Berg, J. J., Lindström, B. R., & Amodio, D. M. (2019). Model-based and model-free social cognition: Investigating the role of habit in social attitude formation and choice. *Frontiers in Psychology*, *10*, Article 2592. https://doi.org/10.3389/fpsyg.2019.02592

Hackel, L. M., Kogon, D., Amodio, D. M., & Wood, W. (2022). Group value learned through interactions with members: A reinforcement learning account. *Journal of Experimental Social Psychology*, *99*, Article 104267. https://doi.org/10.1016/j.jesp.2021.104267

Hackel, L. M., Mende-Siedlecki, P., & Amodio, D. M. (2020). Reinforcement learning in social interaction: The distinguishing role of trait inference. *Journal of Experimental Social Psychology*, *88*, Article 103948. https://doi.org/10.1016/j.jesp.2019.103948

Hackel, L. M., & Zaki, J. (2018). Propagation of economic inequality through reciprocity and reputation. *Psychological Science*, *29*(4), 604–613.

Hafri, A., & Firestone, C. (2021). The perception of relations. *Trends in Cognitive Sciences*, *25*(6), 475–492.

Heider, F. (1958). *The psychology of interpersonal relations*. Wiley.

Holyoak, K. J. (2012). Analogy and relational reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 234–259). Oxford University Press.

Kalkstein, D. A., Bosch, D. A., & Kleiman, T. (2020). The contrast diversity effect: Increasing the diversity of contrast examples increases generalization from a single item. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(2), 296–315.

Kalkstein, D. A., Hackel, L. M., & Trope, Y. (2020). Person-centered cognition: The presence of people in a visual scene promotes relational reasoning. *Journal of Experimental Social Psychology*, *90*, 104009. https://doi.org/10.1016/j.jesp.2020.104009

Kalkstein, D. A., Kleiman, T., Wakslak, C. J., Liberman, N., & Trope, Y. (2016). Social learning across psychological distance. *Journal of Personality and Social Psychology*, *110*(1), 1–19.

Karagoz, A. B., Reagh, Z. M., & Kool, W. (2022). *The construction and use of cognitive maps in model-based control*. PsyArXiv. https://doi.org/10.31234/osf.io/ngqwa

Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychological Science*, *28*(9), 1321–1333.

Kool, W., Gershman, S. J., & Cushman, F. A. (2018). Planning complexity registers as a cost in metacontrol. *Journal of Cognitive Neuroscience*, *30*(10), 1391–1404.

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, *50*(4), 315–353.

Mason, M. F., Magee, J. C., Kuwabara, K., & Nind, L. (2010). Specialization in relational reasoning: The efficiency, accuracy, and neural substrates of social versus nonsocial inferences. *Social Psychological and Personality Science*, *1*(4), 318–326.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61–64.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*(3), 289–316.

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*(2), 185–200.

Otto, A. R., Gershman, S. J., Markman, A. B., & Daw, N. D. (2013). The curse of planning: Dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychological Science*, *24*(5), 751–761.

Park, S. A., Miller, D. S., & Boorman, E. D. (2021). Inferences on a multidimensional social hierarchy use a grid-like code. *Nature Neuroscience*, *24*(9), 1292–1301.

Precup, D., & Sutton, R. S. (1997). Multi-time models for temporally abstract planning. In M. I. Jordan, M. J. Kearns, & S. A. Solla (Eds.), *Proceedings of the 10th International Conference on Neural Information Processing Systems* (pp. 1050–1056). Association for Computing Machinery.

R Core Team. (2016). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing.

Read, S. J. (1987). Constructing causal scenarios: A knowledge structure approach to causal reasoning. *Journal of Personality and Social Psychology*, *52*(2), 288–302.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439.

Shahar, N., Moran, R., Hauser, T. U., Kievit, R. A., McNamee, D., Moutoussis, M., NSPN Consortium, & Dolan, R. J. (2019). Credit assignment to state-independent task representations and its relationship with model-based decision making. *Proceedings of the National Academy of Sciences, USA*, *116*(32), 15871–15876.

Spunt, R. P., & Adolphs, R. (2015). Folk explanations of behavior: A specialized use of a domain-general mechanism. *Psychological Science*, *26*(6), 724–736.

Todorov, A., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*, *83*(5), 1051–1065.

Wang, F., Schoenbaum, G., & Kahnt, T. (2020). Interactions between human orbitofrontal cortex and hippocampus support model-based inference. *PLOS Biology*, *18*(1), Article e3000578. https://doi.org/10.1371/journal.pbio.3000578

Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology*, *47*(2), 237–252.

Wright, J. C., & Murphy, G. L. (1984). The utility of theories in intuitive statistics: The robustness of theory-based judgments. *Journal of Experimental Psychology: General*, *113*(2), 301–322.