Trends in
Cognitive Sciences

CellPress

Opinion

The computational challenge of social learning

Oriel FeldmanHall[1,2,*] and Matthew R. Nassar[2,3]

The complex reward structure of the social world and the uncertainty endemic to social contexts poses a challenge for modeling. For example, during social interactions, the actions of one person influence the internal states of another. These social dependencies make it difficult to formalize social learning problems in a mathematically tractable way. While it is tempting to dispense with these complexities, they are a defining feature of social life. Because the structure of social interactions challenges the simplifying assumptions often made in models, they make an ideal testbed for computational models of cognition. By adopting a framework that embeds existing social knowledge into the model, we can go beyond explaining behaviors in laboratory tasks to explaining those observed in the wild.

Highlights

When interacting with other people, learners face a unique set of challenges.

Standard computational models fail to capture real-world social learning processes.

Improving computational models of social learning will require embracing the complexities of social decision contexts.

Sophisticated social learning models should account for dynamic social rewards, unobservable internal states, unwieldy state–action spaces, and the fact that one person's actions critically influence another's internal state.

The challenge of social learning

As a discipline, social psychology seeks to understand how humans operate in context. Consider deciding whether to collaborate with a colleague, lend money to a friend, or apologize to a loved one. These decisions are not made in isolation. Rather, there is a vast amount of dynamic contextual information that affects a choice, including social norms, past interactions, gossip from peers, the present reaction of a colleague, or the opportunity for future interactions. Through this lens, mapping the social mind attempts to achieve an almost impossible goal: how can we offer a parsimonious, mechanistic account of social learning?

Computational formalisms offer an enticing solution [1]. Well-vetted mathematical models can bring a level of precision and description to fuzzy theories seeking to explain social behaviors in context [2]. For instance, a model, which can come in many flavors and levels of abstraction, might reveal how people learn about one another, whether individual differences affect this learning process, or which contexts promote efficient learning. Indeed, the last decade has witnessed an unprecedented uptick in the use of computational models to test and enumerate the mechanisms governing social behaviors. At first blush, such formalism seems like a promising avenue [3]. However, because the nature of learning changes in fundamental ways when people interact with others [4], standard computational models have struggled to capture the unique learning problems observed in the social world.

Standard computational models of social learning

During social learning, as perhaps with all forms of learning, the problem we are trying to solve is one of inference [5,6]. When interacting with another person, selecting the appropriate next action means figuring out what the other person is currently thinking, feeling, or intending to do. A friendly smile can be the product of a large number of possible unobservable factors (e.g., is he smiling because I just said something funny, because he is happy to be here, or is he thinking about something else entirely unrelated?). Inferring the correct unobservable cause of the smile is essential for deciding the next best action, which in turn influences what the other person will say or do in response.

[1]Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI 02912, USA
[2]Carney Institute for Brain Sciences, Brown University, Providence, RI 02912, USA
[3]Department of Neuroscience, Brown University, Providence, RI 02912, USA

*Correspondence: oriel.feldmanhall@brown.edu (O. FeldmanHall).

The idea that the basic assumptions of reinforcement learning (RL) – where an agent uses reward prediction errors to learn the value associated with specific actions – could be applied to social behaviors has been a boon for social psychology. In their early instantiation, standard computational models of social learning co-opted well-established RL and Bayesian frameworks that were configured to illuminate the mechanisms behind nonsocial learning. These nascent models revealed that, much like classic reward learning, social learning in the laboratory can be explained in terms of prediction errors [7–9]. It turns out that the difference between what a person expects and what they ultimately encounter (the prediction error, δ) can capture how humans learn social value, such as how generous another person is [10], or whether another's desire for punishment influences our own [11]. Other work using basic RL models has established that the rate at which an individual updates their social expectations can be described by a learning rate (α), which appears to be related to stable individual differences, such as racial bias [12] or empathic concern [13].

Standard social RL frameworks typically follow a few key assumptions borrowed from the nonsocial domain [14,15]. For starters, learning occurs incrementally through iterative trial-and-error to estimate the expected value of a choice [8,10,12,16,17]. In addition, reward functions are commonly parameterized by a simple monetary gain or loss [7,11,13], which do not contend with the multifaceted and subjective reward functions that guide most everyday social behavior [18]. Finally, these frameworks typically assume that learning occurs in a context insensitive manner [19–23], such that learning about another's social value, say, their honesty, does not take into account any situational factors that might bias learning (e.g., helping a friend by telling a white lie).

Although the application of basic RL has been successful for explaining many forms of learning in the laboratory, there are reasons to be skeptical of their direct application to social learning problems [24]. For example, in a classic decision-making framework, a reasonable inference after receiving a big payout from a slot machine might be that the mean payout of the machine is high, which should promote future gambling behaviors (Figure 1A). However, the process generating social interactions differs substantially from the one that produces slot machine payouts (Figure 1B). If we were to apply the same logic of interacting with the slot machine to interacting with a person, a joke that elicited a laugh would lead us to infer our partner likes jokes and thus we should simply tell another similar joke. Unless you are 2 years old and charmingly cute, telling similar jokes over and over again would swiftly lead to frowns and the other person may try to slip out the back door to escape.

To deal with this issue, research on social learning has often attempted to minimize this sort of logical discrepancy by restricting the types of social interactions that can occur in laboratory paradigms, or by adding additional parameters to standard models to account for a particular feature of social learning in the most narrow sense (e.g., fitting different learning rates to different partner types in an economic game) [11]. While these types of piecemeal extensions to standard RL models may facilitate better descriptions of data in specific social laboratory paradigms (Box 1), many models built to characterize learning in nonsocial contexts will not generalize to social contexts because the complexities inherent to the social inference problem challenge many of the core assumptions that are made to gain traction when modeling the human mind [25,26] (Figure 1B). To successfully leverage computational models to advance our understanding of how people learn about their social worlds, a more radical reconsideration of the assumptions made in standard models is needed. Rather than recasting a hard problem as a simpler one, we must embark on an epistemological process of judiciously building models and developing tasks that directly grapple with the real-life problem of

Glossary

Internal states: a key component of the social state space is the internal states of the people with whom you interact. This could include internal states regarding what other people know (e.g., 'do they know what I know?'), their emotional state (e.g., 'are they upset?'), or their motives (e.g., 'are they trying to get something from me?'). These internal states are consequential for the outcomes associated with our own actions, limiting the amount of experience that we can accumulate in any particular state.

Joint goals: in the social world, there are many situations where people share a common goal, such as when two more people want to coordinate or collaborate. Examples of this include two colleagues collaborating on a manuscript together, a group of volunteers coordinating to distribute vaccinations or food, a parent and child exhibiting joint attention during teaching, etc.

Latent social structures: these include social norms (e.g., make eye contact when chatting), cultural habits (e.g., stand to the right on British escalators so that others can pass on the left), unobservable psychological motivations (e.g., greed), and so forth.

Reward sources: this could refer to money (as is traditionally done in behavioral economics), or social rewards, such as smiling, being included in a social gathering, physical affection, positive emotions, gossip, etc.

Second-order beliefs: one person's beliefs about another person's beliefs (e.g., 'does he know that I know…that he knows that I know?'). Since internal states often include states of knowledge about socially relevant variables, which in many cases include our own state of knowledge, second and higher order beliefs extend the state space dramatically.

Social priors: prior expectations over socially relevant variables that can be used to constrain interpretations of observed actions (e.g., 'is he acting out of malice, or is there another less devious explanation for his behavior?') and are updated with new observations to provide a dynamic predictive framework for interpreting the ongoing social milieu.

Social reward function: the amount of intrinsic pleasure assigned to all possible achievable outcomes, where such outcomes might be achieved by taking a

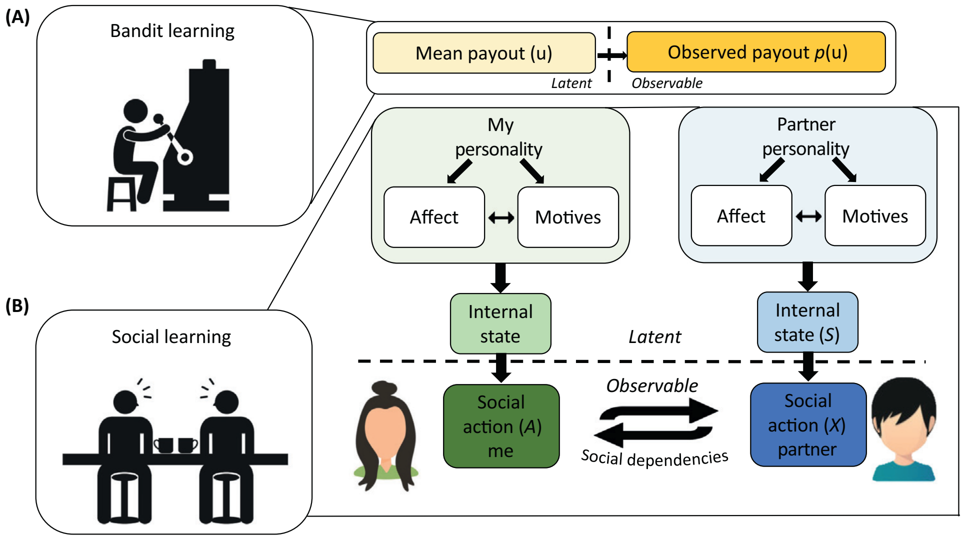


Figure 1. Generative models for the outcome of social and nonsocial decisions. (A) Decisions about slot machines (i.e., bandits) have played a crucial role in shaping our understanding of learning and decision-making in nonsocial domains. (B) While social decisions may share some commonalities with the bandit problem, these situations look much different when we consider the generative model that gives rise to outcomes (e.g., money in the bandit problem or our partner's response in the social domain). In particular, social outcomes depend not only one's own actions (in green: should I tell a joke? Ask on a second date?), but also a large number of variables that collectively make up the internal state of one's partner (in blue). While these social dependencies complicate how we model social learning problems, representing these variables is critical for inferring the partner's underlying state and thus for predicting his response to possible actions that could be taken.

specific social action in a specific state. Such a reward function could be used to define training signals for learning or to assign values for planning.

State–action spaces: the space over possible social actions that could be taken (i.e., telling a joke) and possible social states that they could be taken from (i.e., during a baptism). Two situations should be represented by different states if the actions taken in them differ in their distribution of possible outcomes, in particular, rewards.

interest, which begins with more closely approximating the complex tensions encountered in the real social world.

The complexities of the social world challenge simplifying assumptions made in models

To illustrate the complexities of the social world, let's take the example of going on a first date. As two individuals get to know one another, they share stories that reveal insights about what they value. The date might start out well, where storytelling elicits shared smiles and collective laughter; however, an ill-posed joke leads to a series of frowns, averted eyes, and frequent watch-checking on one side of the table. Imagine the problem each person is trying to solve. Presumably both people aim to maximize their enjoyment and make progress towards possible future rewards; in this case, deciding whether to go on another date. Despite the intuitive simplicity of calculating social rewards, several features of the social world make these calculations difficult (and, in fact, make exact calculations intractable; Figure 2). These features, which do not neatly map on to the assumptions baked into standard computational models, can be grouped into three categories: ubiquitous uncertainty, dynamic rewards in context, and social dependencies, each of which poses a formidable challenge to modeling. That these factors also interact with one another, further compounds the challenge of modeling how learning unfolds in the social world.

Social uncertainty

The first challenge for modeling the social learning problem is the fact that another person's mood, motives, and intentions, which we collectively refer to as their **internal state** (see Glossary), are not observable and are therefore highly uncertain [27,28]. This poses a problem for computing the values of our own potential actions, since they depend on our partners' internal state (Figure 2, purple box) and how our partner will respond to our actions in turn (Figure 2,

Box 1. Expanding social models: what do we gain beyond simple RL?

To concretely illustrate the potential advantages of models that embrace the inherent complexities of social decisions, we consider a trust game where a participant is repeatedly asked how much money to invest in a partner. Once invested, the money is multiplied by some factor and the partner can share some of those increased earnings with the participant. The game can be played iteratively so that the participant can learn about the partner and alter investments accordingly.

A simple RL model

In principle, such a task could be solved through a basic RL algorithm that learns an action value for investing according to the following rule:

$$V_{\{\text{invest}\}} \leftarrow V_{\{\text{invest}\}} + \alpha\left(R - V_{\{\text{invest}\}}\right) \quad [\text{I}]$$

Where R reflects the return on investment (amount returned minus amount invested), α is the learning rate, and V reflects the value of investing. Through trial-and-error (i.e., experiencing monetary returns from the partner), this model could efficiently learn the value of investing in a partner that yields either a stable return on investments or one in which the return on investment slowly shifts over time. With a simple model like this, one could identify, for example, that investing in a predictable partner is more valuable than investing in an unpredictable partner.

What if we build some more realistic social complexity into the problem? Let us consider the possibility that the partner has partially hidden internal states, such as moods. In this case, when the partner is in a happy state, returns tend to be high, and when in an unhappy state, returns tend to be low. If the partner's mood is fairly stable (i.e., transitions between moods are rare), the basic RL algorithm can slowly adjust its investment behavior according to the most recent mood, albeit far from optimally. Other standard off-the-shelf models that deal with discontinuous data, such as hierarchical Gaussian filters [110] or change point models [111], can achieve more efficient learning once the partner's mood has changed, but these models must learn from scratch after each mood transition, rather than transferring previously learned knowledge.

A slightly more complex model

Now consider extending the model to deal directly with the social complexity. Instead of considering the partner as a monolith, let us assume that the partner has multiple internal states that are yoked to how much money is returned to the participant (Figure 2). In this situation, learning requires that the current state of the partner (S) on timestep (t) be inferred first:

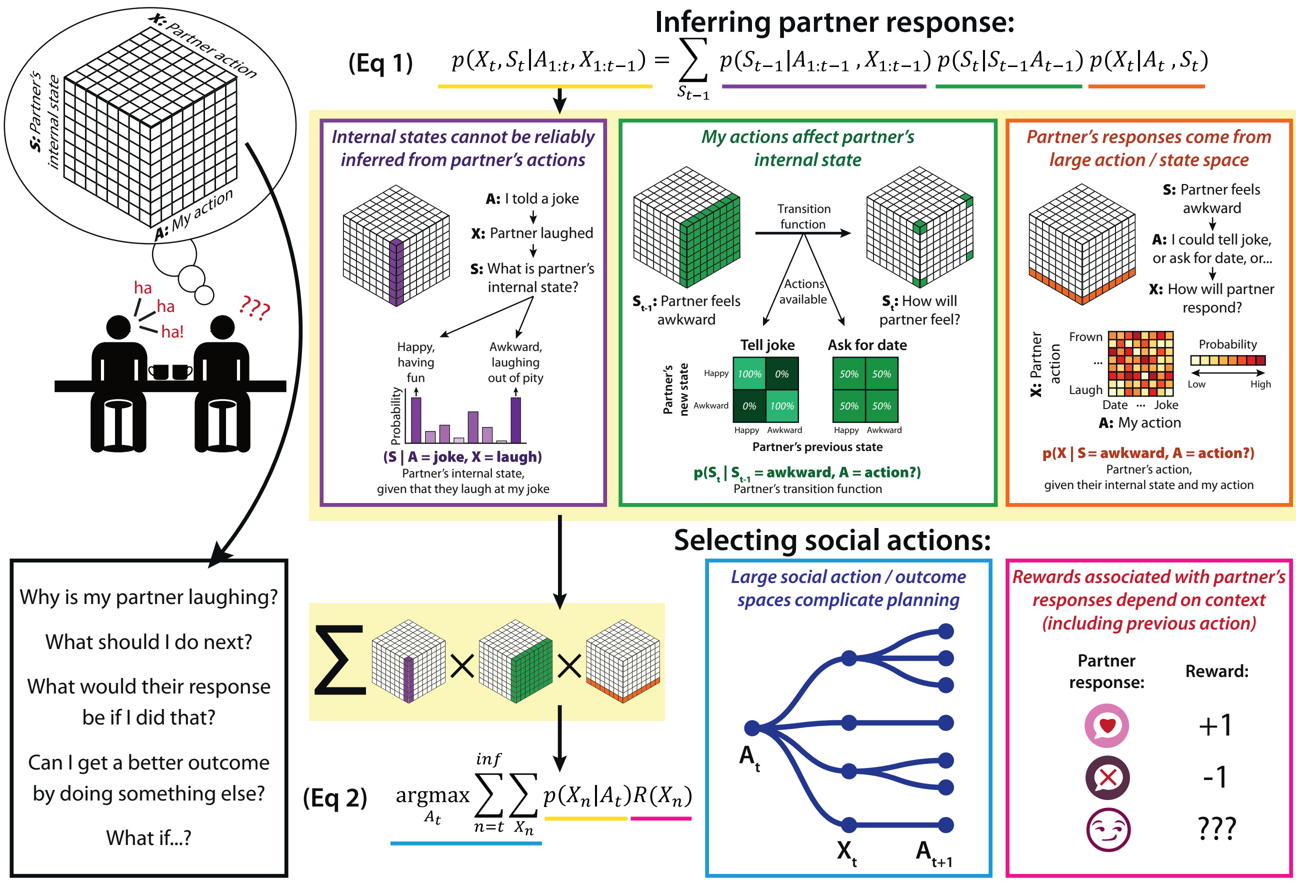
$$p(S_t|S_{t-1}, R_t) = \frac{p(R_t|\, S_t)\, p(S_t \mid S_{t-1}) p(S_{t-1})}{p(R_t)} \quad [\text{II}]$$

Based on the likelihood of that state yielding the observed return (R) and the probability of a transition from the previous state (mood) to the current one $p(S_t \mid S_{t-1})$. Like the simple RL algorithm earlier, a participant must learn the return rate associated with a partner being in a given mood, which could be done with Bayes rule or through an RL approximation:

$$\hat{R}_{S_t=s} = \hat{R}_{S_t=s} + \alpha\left(R_t - \hat{R}_{S_t=s}\right) \quad [\text{III}]$$

Where $\hat{R}_{S_t=s}$ is the return expected in the current state (analogous to $V_{\{\text{invest}\}}$ in the simple model) and R_t is the actual return experienced. The key difference between this model and the simple RL model is that the more complex model will be slow to adjust after state transitions that occur early in the game, as expected returns given a specific state are still unclear. Once the model learns that there are two common states that have high and low average returns, respectively, it only needs to see one trial after a transition to quickly infer that the partner's mood has changed and expectations about future returns are adjusted accordingly. The fact that this slightly more complex model can capture behaviors that are better matched to this particular task and provide an intuitive explanation that participants learn to recognize (and value) the moods of their partner, yields a tool and result that are more likely to generalize to real-world settings.

Nonetheless, even this toy problem and model are highly simplified with respect to social learning in the wild and will thus be unable to capture other aspects of social behavior observed beyond the conditions of this particular task. For example, recognizing that a partner is in a bad mood might lead someone to try to cheer the partner up, which could be viewed from the utilitarian perspective as selecting actions that are likely to promote latent state transitions that yield favorable future social interactions. Although this social learning problem could be modeled in the same framework by expanding the transition function $p(S_t \mid S_{t-1})$ to include the actions of the investor $p(S_t \mid S_{t-1}, A_t)$, testing such a model in the laboratory requires a more complex task design (e.g., action selection affects the state transition structure), which reduces experimental control (as participants now control how they move through state space). We do not advocate giving up experimental control completely, but instead, leveraging task designs and models that systematically explore specific features of social learning, while ensuring sufficient experimental control to adequately characterize these features.



Trends in Cognitive Sciences

Figure 2. Mathematics of social learning. Left: social situations, such as a first date, require considering a host of possibilities to better understand why a partner is behaving a certain way. These considerations are useful in that they might allow us to better predict how someone would respond if we chose a particular action, thereby promoting social actions with desirable outcomes. Despite the ubiquity of these social situations, and our ability to navigate them effectively, modeling these dynamics poses a number of mathematical challenges. Top: social learning requires inferring the internal state of the partner (S) and their future action (X_t) based on previous observations of their responses ($X_{1:t-1}$) to our own actions ($A_{1:t-1}$). Such inferences are highly uncertain, as different internal states could yield similar behavioral responses (purple box). Furthermore, learning to predict a partner's response to our own actions requires considering how our partner's internal state is changing, which depends critically on our own actions: saying something boring is unlikely to change the internal state of a partner, but proposing a second date might do so, one way or the other (green box). A final issue is that the combinatorial size of all potential actions and partner internal states means that most available actions have never been tried in the current state (orange box). Only by considering (multiplying) all of these factors, and integrating over internal states, can we determine our partner's most likely response to our candidate actions (underlined in yellow). Bottom: within a normative framework, our own social actions (A) should be selected according to their ability to maximize long run rewards (R). Doing so requires computing the probability distribution over our partner's response to each of our candidate actions on the current timestep ($p(X_n|A_t)$; yellow shading) as described earlier, then multiplying it by the reward it would yield ($R(X_n)$), and doing so recursively through future timesteps to evaluate the sequence of actions that would yield the largest long run returns (Equation 2). Inferring a partner's response to our actions ($p(X|A)$) comes with a number of challenges. Planning is also particularly difficult in social domains because of the size of the space of possible actions and associated responses (blue box). An additional issue is that the reward associated with any given outcome (e.g., your partner's response) is highly contextual [e.g., the reward associated with a partner laughing may be high if it follows a joke, but low if it follows a proposal for a second date (pink box)].

orange box). People are constantly trying to figure out what others know or believe, which is only made more challenging by the fact that people are not static, reliable, or predictable and are subject to the influence of emotion and past experience [29]. Moreover, each person we encounter is unique and has a different set of ever-changing intentions, beliefs, and moods. The combination of stochastic (i.e., no two people are the same) and nonstationary (i.e., each person's internal state changes; Figure 2, green box) uncertainty, only adds an additional hurdle when trying to formally model how people infer what another's motivations are [30].

Uncertainty also arises due to the breadth of the action space. Not having access to the thoughts and beliefs of others means there is no fixed action space where the same action, or even set of actions, leads to a certain likelihood of reward (in this case, a smile or second date). The difficulty with parameterizing what the action space even ought to be when there are people dynamically interacting creates a dimensionality problem. Any conceived action (e.g., tell a joke, ask on a second date) and state (e.g., laughter, boredom) is possible with one person and when there are two or more interacting individuals, the **state–action space** explodes (Figure 2, yellow shaded box). The breadth of this space not only makes the values associated with each action more uncertain, but it also increases the difficulty of the planning problem (Figure 2, blue box). Despite the uncertainty and planning difficulties produced by social state–action spaces, people routinely display social behaviors that demonstrate long-run planning, such as asking a friend to call at a certain time in case the date is going poorly. That humans can achieve such sophisticated learning provides a challenge to models in formalizing how this is done.

Rewards in context

The second issue is how models should account for dynamic rewards with multiple interacting people. Rewards are critical for computing the expected value of social actions (Figure 2, pink box), yet social rewards are highly contextual and critically depend on the wants, needs, and motivations of the individuals interacting, as well as the collective dynamic between them [31,32]. In our example, halfway through the date, one person may be looking for a second date while the other is plotting the best excuse to leave early. Even within an individual, a partner's laugh might be perceived as rewarding in one context (after telling a joke) but may have negative implications in another (following a request for a second date), revealing just how sensitive social rewards are to the situation. Because rewards can be inconsistent within a particular person and are not necessarily shared across people, models that assume a singular, fixed reward function are likely to capture only the most simplified social learning problems.

Figuring out the proper reward function becomes even more challenging once you take into account uncertainty and, in particular, self-referential, **second-order beliefs** about what others think you know [33]. When one person assumes the other has a different reward function than they actually do (e.g., misreading the social cues that the date is metaphorically over), there is a misalignment between the structure of interacting social minds and the algorithms meant to assess the computations governing that interaction. Because reward is partially hidden and differently parameterized for each person, there is the ever persistent challenge of formalizing **joint goals** (e.g., coordinating a second date or agreeing not to go out again) [34].

Social dependencies

As if this were all not challenging enough, social inference problems become even thornier when we consider that one person's actions affect another's internal state. A person's choice to say something (or to not say something), express a certain emotion, or make a joke, can swiftly alter another's feelings and motivations (Figure 2, green box). This more complicated inference problem adds challenges, but also an opportunity, for models: rather than trying to acquire immediate rewards, effective planning could allow for the selection of actions that influence a partner's internal state, which in turn would help reap rewarding outcomes generated in that state. This sort of behavior is ubiquitous in social interactions (e.g., paying a compliment before asking a favor) [35], but scaling up models to achieve this behavior is not trivial. For every person you add into the mix (e.g., group dynamics), these issues blossom, as the actions of each person can affect the internal state of every other member in the group.

These multiple layers of recursion and uncertainty propagation endemic to the social world make it difficult to formalize complex social interactions in a mathematically tractable way. However, the fact that the number of factors known to bias social cognition is matched by the sheer number of verbal theories that abound [36] only highlights the pressing need to more precisely characterize the mechanics of the social mind. The problem is that reducing the theoretical landscape through a superficial model-fitting enterprise informed by simple, borrowed models (e.g., focusing on model estimation with little attention to model criticism [37]), has the potential to lead us astray [38]. When simple models are fit to complex behavior, parameter estimates can be misleading [25,39,40] and will likely fail to generalize beyond a particular experimental paradigm [24], a point that has recently been noted with respect to the learning rate in RL models [41]. The danger is even greater when simple models constrain experimental paradigms, since it has the potential to divert resources to solving a tractable problem that bears little resemblance to the realities of the social world [42,43]. In either case, the resulting mechanistic conclusions can be misleading and are thus unlikely to illuminate much that is useful about *how* humans actually navigate their social worlds. So how do we create models that emulate the unique dynamics of the social world?

A theoretical solution: generate models inspired by social psychology (rather than vice versa)

We believe that modeling social interactions provides an ideal testbed for computational models of cognition, *precisely because* the complexities of the social world challenge simplifying assumptions often made in off-the-shelf models. By using social cognition as the testing ground, we can redefine how models are created and deployed, which means they can more closely approximate the tensions encountered in the real world. In many cases, this means casting off some of the tried-and-true frameworks that have been instrumental for probing other types of cognitive processes. We can replace those models with hypothesis-driven and well-specified computational frameworks that reflect the structure of the social environment (e.g., the other people present, the broad social norms, the local social habits), which will be able to provide necessary insights about the inner architecture of the social mind. Moreover, to successfully capture the complexities of the social world, researchers should start by staking the question on psychological grounds. The range of cognitive processes in question should influence how a paradigm is constructed (ideally by generating one that closely mimics the structure of the social context) and how a computational model is created (ideally one that interrogates multiple mechanisms that could be at play). If theories of social learning are formulated thoughtfully in computational terms, then verbal models can be replaced with those that are mathematically tractable, which enables psychological assumptions and mechanisms to be clearly tested. To harness the power of computational modeling when trying to understand interacting human minds, we offer three concrete strategies, each of which help to mitigate the three challenges detailed earlier.

Creating complex social learning paradigms

Given that models and paradigms are intimately linked, one way to produce a model that can better reflect the social mind is to construct complex paradigms that are imbued with some of the uncertainties present in the social world. This requires any model that provides a sufficient explanation of behavior to cope with these complexities. Rather than relying on artificial, constricted analogs of how state–action space is organized when people interact, more complex paradigms enable the existence, and therefore measurement, of multiple state spaces and social representations [44–47]. The problem of stripping a task of its complexity and, by extension, limiting the amount of social uncertainty present in the testbed, is brought into stark relief when we consider the problem of generalization (i.e., when to take previously learned rules and apply them to novel contexts). Because no two situations are alike in the social world, successfully generalizing is of paramount importance [48,49]. Consider the situation in which you repeatedly ask your colleague for a favor. In some cases, he readily agrees, but in others he fails to offer any help. If you could figure out why he only occasionally helps it would save you a lot of time and energy, since you could use this information to guide whether future requests should be made. Fortunately, just beneath the surface of every interaction lie stable social structures, which include unobservable psychological motivations [50] and shared latent social norms [51]. Although the inference problem in the social world is hard, people can make use of these latent dependencies to minimize uncertainty and serve their own goals – otherwise, inference in the space of real-world behavior would be completely intractable without strong inductive biases [52,53].

Building paradigms that capitalize on these **latent social structures** can help researchers understand how generalization mechanisms are deployed during social interactions rife with uncertainty [54,55]. For example, in our own work, we constructed a paradigm to test how humans figure out the unobservable motivations of another person (e.g., a player's greed, envy, risk aversion, etc.) from their observable decisions (e.g., cooperating or defecting with opponents [44]). Players and opponents interacted in many different economic games, where the monetary payoffs continually changed, reflecting the evolving, multifaceted pressures of the social world. Despite different contexts and shifting tensions in each specific game, a player's motivations remain relatively stable. To make it concrete, a player who is motivated by greed will always try to maximize monetary payoff, which means cooperating in some games but defecting in others, depending on the local payoff matrix. In contrast, a player motivated by risk aversion will try to minimize potential losses, which will yield a different structured choice pattern across games. It is only by observing behaviors across a number of contexts, each with different tensions, that an individual can infer the latent structure (i.e., motivation) associated with a player's choice patterns. The better people are at inferring another player's hidden motivation, the more successfully they use this information in novel contexts to competitively outwit that player. We used model comparison to test competing psychological accounts of *how* humans might learn about another's unobservable motivations. A model that captured the link between discrete choice patterns and abstracted latent states outperformed models optimized for granular learning (i.e., learning the value of each discrete stimulus, which would be both costly and useless for generalizing to new situations). Only by leveraging a formal model that capitalized on uncertainties found in the social world could we show that correctly inferring the motives of another person rests in the ability to combine prior knowledge with specific learning strategies.

Building multifaceted social reward functions

Models of social learning should also account for the fact that every person has a different reward function and this **social reward function** can change depending on the context or group dynamic. One particular model class, called inverse RL, provides a framework through which reward functions can be inferred, rather than hard coded, thus bridging the disconnect between traditional RL algorithms and actual social learning [56,57]. In inverse RL, observed behaviors are used to infer the unknown structure of a person's preferences [57,58], which enables us to learn from others even when they do not share our particular set of predilections [59]. In one clever experiment, the researchers set up the enduring lunch problem many of us are acquainted with – determining how much time to spend walking around in search of our desired food truck [60]. Subjects were asked to predict another's food preferences after watching that person navigate through the local environment in search of the perfect lunch. The researchers found that beliefs about another's preferences incorporated three causal factors: what an agent perceives, the agent's prior beliefs, and the causes of the agent's actions. Just like in the real world, the model enabled inference to occur in a multitude of ways. We can make forward inferences about what others believe from their percepts, we can backwardly infer a certain desire based on actions, or some combination of both. The most sophisticated model incorporated mentalistic inference and therefore was able to capture social learning far better than the other more impoverished models that did not formally model the mentalizing processes. Moreover, the winning model was successful at capturing inferences about the agent's beliefs across a range of situations. Generative models that incorporate evidence from social and developmental science, detailing the depth and flexibility by which people think about the hidden states of others, have helped better characterize the precise computations governing the high-level capacity to mentalize about others [61].

Researchers will be able to develop useful computational models that better account for the complexities of the social world if they additionally allow learning to occur from multiple and evolving **reward sources** [62]. At the most basic level, value functions in social models should not be optimized to learn from many, exact repetitions (a behavioral pattern that rarely occurs in real life), but instead should be optimized to capture both incremental and dramatically shifting rewards [46]. A computational model that reflects the reality of our social worlds should also ideally have a reward function that accounts for the difference between rewards accrued for the self and those accrued by others [63,64], which would enable rewards to be computed in group settings [45]. Furthermore, reward functions sensitive to context can reveal nuance in how people construct and represent social value. Rather than learning about an individual's global social value (e.g., generosity), people in the real world routinely attend to situational factors that can bolster or mitigate value representation (e.g., only making donations in the presence of others).

Perhaps the most challenging problem is creating a reward function that can capture the broad range of rewarding stimuli in the social world (smiles, thoughtful emails, party invitations, marriage, etc.) all of which can be diagnostic when figuring out how to behave with others. Since emotions are known to provide relevant feedback to foster useful action tendencies [65,66], they are well situated to translate external rewards into an internal value signal [67–70]. For example, emotions may provide both real (e.g., joy) and fictive (e.g., regret) value signals that can fine-tune learning [71]. Applying a formal structure for how emotion biases learning [72–74], perhaps by associating an affective experience with prediction errors [70], can simplify and organize how disparate social rewards feed into a biologically plausible reward function.

Leverage prior research to inform the model

To account for the fact that in the social world one person's actions affect another's internal state, researchers should turn to the trove of empirical work conducted in social psychology and behavioral economics that has detailed the nature of these social dependencies [75–77]. Rather than rediscovering information from scratch, embracing existing knowledge about attribution [78], attitudes [50], social influence [79], group dynamics [80], and social preferences [81–84], and embedding this knowledge in computational frameworks, promises more well-specified questions and models [85]. This strategy can be implemented at various levels of abstraction, including how we determine what is being represented (e.g., which boxes get included in Figure 1 or the values we assume for each box) and the architecture of those representations (e.g., how the boxes connect). For example, category-based expectations of how others should behave [86] shape the moral obligations we place on others [87] and can swiftly bias the impressions we form [88]. Even altering just a few of the assumptions in a more traditional RL framework to incorporate this social knowledge has been successful in describing how humans navigate the social learning problem. For instance, a negative first impression can be so powerful that people may ignore subsequent positive social information, effectively abolishing a learning effect [89]. On the flip side, informing the model of an individual's **social priors** or beliefs about the world [16,90] can speed the learning rate [77,91] and these models fair far better in describing how people actually learn in strategic situations [61].

Well-informed social computational models become particularly useful when they are built to adjudicate between existing social psychological theories that offer opposing learning accounts. Humans learn in a variety of ways, including by observing the behaviors of others. Observational learning can occur through emulation, in which individuals infer the other agent's goals and intentions, or through imitation, in which individuals mimic the actions of other agents. Thus, emulation and imitation formalize two assumptions about the depth of inference problem being solved, in particular, whether the inference includes another's internal state. Although theoretical arguments suggest that culture is propagated through pure imitation, conflicting evidence reveals how little we know about when emulation or imitation is used as an efficient learning strategy [92,93]. By incorporating prior theoretical and empirical research from social and developmental psychology into a formal computational framework, we can figure out when these learning strategies might dominate. Researchers recently examined whether uncertainty modulates the degree to which an individual switches between imitation or emulation [94], hypothesizing that if a teacher's observed actions become too stochastic, emulation should be favored. If ascribing the link between choice and goal inference during emulation becomes too difficult, however, then an imitation strategy should be favored. Through model comparison, the researchers identified that observational learning is actually a hybrid of both strategies, where the degree of uncertainty acts as a dial that dictates which of the two strategies is deployed in the moment. By embracing the complexities of a real-world problem and formalizing them within the model, the researchers were able to successfully test simplifying assumptions embedded in standard models to reveal that emulation is only favored once the learning signal becomes sufficiently reliable, elsewise imitation is the default strategy.

At minimum, to avoid psychologically implausible mechanisms, social models should not operate on generic representational or architectural assumptions [95]. Take the example of exploration and how people collect information to optimize their future behavior. Despite accumulating evidence that people are selective about when and how they explore [96], many RL models instantiate algorithms that sample actions at random [15,97]. Applying random exploration algorithms to simplified task designs can corrupt measures of exploration, making the algorithms susceptible to biases and, at the end of the day, hard to interpret [24,25,97]. In the social world, one way to explore is to observe others. Rather than indiscriminately learn from anyone (which would bear out generations of fools), we use information about experience and expertise, which helps when deciding who to learn from [98–102]. Research reveals that people wisely neglect the naïf and attend to those who have experience [103]. By explicitly outlining this social knowledge in the model, we have gained a flourish of new insights into the computational representations of the social mind [102,104–107].

Concluding remarks

By adopting a framework inspired by our existing knowledge about social, developmental, and cognitive learning processes, we can go beyond explaining behaviors in laboratory tasks to explaining those that occur between multiple interacting people in the wild. Even though navigating through the social world is an enduring challenge, people routinely display behaviors demonstrating that they can successfully learn about, and from, other people. To produce a model that reflects these real-world successes, we need to carefully consider the impact of simplifying assumptions, construct paradigms that reveal the complexity of our social representations, allow learning to occur from multiple reward sources, and formulate social learning theories in precise,

Outstanding questions

What is the representational space that we use to characterize the internal state of our social partners? How do the dimensions of this space depend on the extent of our interaction with this partner (e.g., waitress at a restaurant versus our date)?

How do people simplify the complexity of the intractable inference problem of social learning? What information is lost through such simplifications and what pathological behaviors might emerge as a consequence?

Do people use complex learning strategies when first working through a novel inference problem, but transfer to simpler, more biologically plausible strategies once some knowledge is acquired? How can we create models to reflect multifaceted learning strategies?

Is there a fixed mapping between social behavior from impoverished social tasks commonly fit with computational models and the complex behaviors observed in the wild? If so, which dimensions of real-world behavior map onto the things that are most commonly measured in the laboratory and which do not?

How can experiments be designed to maximize the identifiability and construct validity of the reward function?

mathematically tractable terms. There is no doubt a tension between the complexity of a social model and its tractability for standard uses and we are not suggesting that a 'model of everything' needs to come before a 'model of anything'. Instead, we advocate for the construction of tasks and models that embrace the complexity necessary to answer important questions in social learning, while also employing assumptions that limit complexities unrelated to the question of interest. As others have said before us, we see mathematical precision as potentially beneficial, but not a panacea, especially when the mathematical tools are ill-equipped to capture the complexity of the social world. Indeed, simple models can tempt us to substitute a mathematically tractable problem for the one that we really want to answer – a peril of statistical modeling that statistician George Box referred to as 'mathematistry' [42,43].

However, with the proper tools in hand, there are many questions that can be readily tackled (see Outstanding questions). Together, employing these practices will enable the social learning problem to be bounded and will buffer against ill-specified models that can lead to misinterpretations about what gives rise to social behaviors. Of course these principles should be adopted alongside good model practices, for example, employing model comparison, doubting the winning model's assumptions, testing whether the model generalizes to new contexts, creating generative, rather than descriptive models, and the list goes on [108,109]. Our hope is to raise awareness about the issues that arise when applying computational models to social interactions and provide a theoretical scaffold that can bridge social cognition and computational psychology. This can provide a promising foundation for measurable progress for years to come.

Acknowledgments

We thank Michael J. Frank for thoughtful comments during spirited conversations and insightful feedback on earlier drafts of the manuscript. We also thank Jae-Young Son and Tiantian Li for their help in depicting the social learning problem visually.

Declaration of interests

No interests are declared.

References

1. Cushman, F. and Gershman, S. (2019) Editors' introduction: computational approaches to social cognition. *Top. Cogn. Sci.* 11, 281–298
2. Lee, D. (2013) Decision making: from neuroscience to psychiatry. *Neuron* 78, 233–248
3. Olsson, A. *et al.* (2020) The neural and computational systems of social learning. *Nat. Rev. Neurosci.* 21, 197–212
4. Wheatley, T. *et al.* (2019) Beyond the isolated brain: the promise and challenge of interacting minds. *Neuron* 103, 186–188
5. Malle, B.F. and Holbrook, J. (2012) Is there a hierarchy of social inferences? The likelihood and speed of inferring intentionality, mind, and personality. *J. Pers. Soc. Psychol.* 102, 661–684
6. Hastie, R. (1983) Social inference. *Annu. Rev. Psychol.* 34, 511–542
7. Burke, C.J. *et al.* (2010) Neural mechanisms of observational learning. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14431–14436
8. Behrens, T.E. *et al.* (2008) Associative learning of social value. *Nature* 456, 245–249
9. van den Bos, W. *et al.* (2013) Neural correlates of reinforcement learning and social preferences in competitive bidding. *J. Neurosci.* 33, 2137–2146
10. Hackel, L.M. *et al.* (2015) Instrumental learning of traits versus rewards: dissociable neural correlates and effects on choice. *Nat. Neurosci.* 18, 1233
11. FeldmanHall, O. *et al.* (2018) Learning moral values: another's desire to punish enhances one's own punitive behavior. *J. Exp. Psychol. Gen.* 147, 1211–1224
12. Lindstrom, B. *et al.* (2014) Racial bias shapes social reinforcement learning. *Psychol. Sci.* 25, 711–719
13. Lockwood, P.L. *et al.* (2016) Neurocomputational mechanisms of prosocial learning and links to empathy. *Proc. Natl. Acad. Sci. U. S. A.* 113, 9763–9768
14. Lockwood, P.L. and Klein-Flugge, M.C. (2021) Computational modelling of social cognition and behaviour-a reinforcement learning primer. *Soc. Cogn. Affect. Neurosci.* 16, 761–771
15. Sutton, R.S. and Barto, A.G. (1998) *Reinforcement Learning: An Introduction*, MIT Press
16. Zhu, L. *et al.* (2012) Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proc. Natl. Acad. Sci. U. S. A.* 109, 1419–1424
17. Zaki, J. *et al.* (2016) Social cognition as reinforcement learning: feedback modulates emotion inference. *J. Cogn. Neurosci.* 28, 1270–1282
18. Behrens, T.E.J. *et al.* (2009) The computation of social behavior. *Science* 324, 1160–1164
19. Hein, G. *et al.* (2016) How learning shapes the empathic brain. *Proc. Natl. Acad. Sci. U. S. A.* 113, 80–85
20. Jones, R.M. *et al.* (2011) Behavioral and neural properties of social reinforcement learning. *J. Neurosci.* 31, 13039–13045
21. Meffert, H. *et al.* (2015) Prediction errors to emotional expressions: the roles of the amygdala in social referencing. *Soc. Cogn. Affect. Neurosci.* 10, 537–544
22. Suzuki, S. *et al.* (2012) Learning to simulate others' decisions. *Neuron* 74, 1125–1137
23. Apps, M.A. and Ramnani, N. (2014) The anterior cingulate gyrus signals the net value of others' rewards. *J. Neurosci.* 34, 6190–6200
24. Eckstein, M.K. and Collins, A.G.E. (2021) What do reinforcement learning models measure? Interpreting model

parameters in cognition and neuroscience. *Curr. Opin. Behav. Sci.* 41, 128–137
25. Nassar, M.R. and Gold, J.I. (2013) A healthy fear of the unknown: perspectives on the interpretation of parameter fits from computational models in neuroscience. *PLoS Comput. Biol.* 9, e1003015
26. van den Bos, W. *et al.* (2018) Computational neuroscience across the lifespan: promises and pitfalls. *Dev. Cogn. Neurosci.* 33, 42–53
27. Jern, A. and Kemp, C. (2015) A decision network account of reasoning about other people's choices. *Cognition* 142, 12–38
28. Jara-Ettinger, J. (2019) Theory of mind as inverse reinforcement learning. *Curr. Opin. Behav. Sci.* 29, 105–110
29. Ross, L. and Nisbett, R.E. (1991) *The Person and the Situation: Perspectives of Social Psychology*, McGraw-Hill
30. FeldmanHall, O. and Shenhav, A. (2019) Resolving uncertainty in a social world. *Nat. Hum. Behav.* 3, 426–435
31. Ho, M.K. *et al.* (2017) Social is special: a normative framework for teaching with and learning from evaluative feedback. *Cognition* 167, 91–106
32. Juechems, K. and Summerfield, C. (2019) Where does value come from? *Trends Cogn. Sci.* 23, 836–850
33. Asaba, M. and Gweon, H. (2021) Young children rationally revise and maintain what others think of them. *Nature Communications*. *PsyArXiv* Published online December 19, 2019. http://doi.org/10.31234/osf.io/yxhv5
34. Mastroianni, A.M. *et al.* (2021) Do conversations end when people want them to? *Proc. Natl. Acad. Sci. U. S. A.* 118, e2011809118
35. Cialdini, R.B. (2001) The science of persuasion. *Sci. Am.* 284, 76–81
36. Ferguson, C.J. and Heene, M. (2012) A vast graveyard of undead theories: publication bias and psychological science's aversion to the null. *Perspect. Psychol. Sci.* 7, 555–561
37. Box, G.E.P. (1980) Sampling and Bayes' inference in scientific modelling and robustness. *J. R. Stat. Soc. Ser. A (General)* 143, 383–404
38. Nassar, M.R. and Frank, M.J. (2016) Taming the beast: extracting generalizable knowledge from computational models of cognition. *Curr. Opin. Behav. Sci.* 11, 49–54
39. Zhang, L. *et al.* (2020) Using reinforcement learning models in social neuroscience: frameworks, pitfalls and suggestions of best practices. *Soc. Cogn. Affect. Neurosci.* 15, 695–707
40. Collins, A.G.E. *et al.* (2014) Working memory contributions to reinforcement learning impairments in schizophrenia. *J. Neurosci.* 34, 13747–13756
41. Eckstein, M.K. *et al.* (2021) Learning rates are not all the same: the interpretation of computational model parameters depends on the context. *bioRxiv* Published online May 28, 2021. https://doi.org/10.1101/2021.05.28.446162
42. Box, G.E.P. (1976) Science and statistics. *J. Am. Stat. Assoc.* 71, 791–799
43. Navarro, D.J. (2019) Between the devil and the deep blue sea: tensions between scientific judgement and statistical model selection. *Comput. Brain Behav.* 2, 28–34
44. van Baar, J. *et al.* (2021) Latent motives guide structure learning during adaptive social choice. *bioRxiv* Published online June 8, 2020. https://doi.org/10.1101/2020.06.06.137893
45. Park, S.A. *et al.* (2019) Neural computations underlying strategic social decision-making in groups. *Nat. Commun.* 10, 5287
46. Lamba, A. *et al.* (2020) Anxiety impedes adaptive social learning under uncertainty. *Psychol. Sci.* 31, 592–603
47. Steixner-Kumar, S. *et al.* (2020) Humans depart from optimal computational models of socially interactive decision-making under partial information. *PsyArXiv* Published online November 3, 2020. http://doi.org/10.31234/osf.io/rcq25
48. FeldmanHall, O. and Dunsmoor, J.E. (2018) Viewing adaptive social choice through the lens of associative learning. *Perspect. Psychol. Sci.* 14, 175–196
49. FeldmanHall, O. *et al.* (2018) Stimulus generalization as a mechanism for learning to trust. *Proc. Natl. Acad. Sci. U. S. A.* 115, E1690–E1697
50. Heider, F. (1946) Attitudes and cognitive organization. *J. Psychol.* 21, 107–112
51. Ostrom, E. (2000) Collective action and the evolution of social norms. *J. Econ. Perspect.* 14, 137–158
52. Hessel, M. *et al.* (2019) *On inductive biases in deep reinforcement learning*. arXiv.1907.02908
53. Griffiths, T.L. *et al.* (2010) Probabilistic models of cognition: exploring representations and inductive biases. *Trends Cogn. Sci.* 14, 357–364
54. Lau, T. *et al.* (2018) Discovering social groups via latent structure learning. *J. Exp. Psychol. General* 147, 1881–1891
55. Gershman, S.J. *et al.* (2017) Learning the structure of social influence. *Cogn. Sci.* 41, 545–575
56. Ng, A.Y. and Russell, S.J. (2000) Algorithms for inverse reinforcement learning. In *Proceedings of the Seventeenth International Conference on Machine Learning*, pp. 663–670, Morgan Kaufmann Publishers
57. Baker, C.L. *et al.* (2009) Action understanding as inverse planning. *Cognition* 113, 329–349
58. Collette, S. *et al.* (2017) Neural computations underlying inverse reinforcement learning in the human brain. *Elife* 6, e29718
59. Jern, A. *et al.* (2017) People learn other people's preferences through inverse decision-making. *Cognition* 168, 46–64
60. Baker, C.L. *et al.* (2017) Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nat. Hum. Behav.* 1, 64
61. Hampton, A.N. *et al.* (2008) Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. U. S. A.* 105, 6741–6746
62. Fu, J. *et al.* (2018) *Learning robust rewards with adversarial inverse reinforcement learning*. arXiv.1710.11248
63. Yoshida, W. *et al.* (2008) Game theory of mind. *PLoS Comput. Biol.* 4, e1000254
64. Coricelli, G. and Nagel, R. (2009) Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 106, 9163–9168
65. Dunsmoor, J.E. *et al.* (2015) Emotional learning selectively and retroactively strengthens memories for related events. *Nature* 520, 345
66. Phelps, E.A. *et al.* (2014) Emotion and decision making: multiple modulatory neural circuits. *Annu. Rev. Neurosci.* 37, 263–287
67. Wu, Y. and Gweon, H. (2021) Preschool-aged children jointly consider others' emotional expressions and prior knowledge to decide when to explore. *Child Dev.* 92, 862–870
68. Rutledge, R.B. *et al.* (2014) A computational and neural model of momentary subjective well-being. *Proc. Natl. Acad. Sci. U. S. A.* 111, 12252–12257
69. Eldar, E. and Niv, Y. (2015) Interaction between emotional state and learning underlies mood instability. *Nat. Commun.* 6, 6149
70. Heffner, J. *et al.* (2021) Emotion prediction errors guide socially adaptive behavior. *Nat. Hum. Behav.* Published online June 3, 2021. http://doi.org/10.31234/osf.io/azeyk
71. FeldmanHall, O. and Chang, L.J. (2019) Social learning: emotions aid in optimizing goal-directed social behavior. In *Goal-Directed Decision-Making: Computations and Neural Circuits* (Morris, R. *et al.*, eds), pp. 309–330, Academic Press
72. Ong, D.C. *et al.* (2019) Computational models of emotion inference in theory of mind: a review and roadmap. *Top. Cogn. Sci.* 11, 338–357
73. Eldar, E. *et al.* (2016) Mood as representation of momentum. *Trends Cogn. Sci.* 20, 15–24
74. Wu, Y. *et al.* (2020) Emotion as information in early social learning. *PsyArXiv* Published online October 26, 2020. http://doi.org/10.31234/osf.io/fz857
75. Heider, F. (1958) *The Psychology of Interpersonal Relations*, John Wiley & Sons
76. Von Neumann, J. and Morgenstern, O. (1944) *Theory of Games and Economic Behavior*, Princeton University Press
77. Camerer, C. (2003) *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press
78. Kelley, H.H. and Michela, J.L. (1980) Attribution theory and research. *Annu. Rev. Psychol.* 31, 457–501
79. Turner, J.C. (1991) *Social Influence*, Thomson Brooks/Cole Publishing

80. Chen, Y. and Li, S.X. (2009) Group identity and social preferences. *Am. Econ. Rev.* 99, 431–457
81. Fehr, E. and Schmidt, K.M. (1999) A theory of fairness, competition, and cooperation. *Q. J. Econ.* 114, 817–868
82. Charness, G. and Rabin, M. (2002) Understanding social preferences with simple tests. *Q. J. Econ.* 117, 817–869
83. Levitt, S.D. and List, J.A. (2007) What do laboratory experiments measuring social preferences reveal about the real world? *J. Econ. Perspect.* 21, 153–174
84. Fischbacher, U. and Gachter, S. (2010) Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *Am. Econ. Rev.* 100, 541–556
85. Hackel, L.M. and Amodio, D.M. (2018) Computational neuroscience approaches to social cognition. *Curr. Opin. Psychol.* 24, 92–97
86. Gopnik, A. and Wellman, H.M. (2012) Reconstructing constructivism: causal models, Bayesian learning mechanisms, and the theory. *Psychol. Bull.* 138, 1085–1108
87. Rhodes, M. and Chalik, L. (2013) Social categories as markers of intrinsic interpersonal obligations. *Psychol. Sci.* 24, 999–1006
88. Fiske, S.T. and Neuberg, S.L. (1990) A continuum of impression-formation, from category-based to individuating processes - influences of information and motivation on attention and interpretation. *Adv. Exp. Soc. Psychol.* 23, 1–74
89. Delgado, M.R. *et al.* (2005) Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618
90. Diaconescu, A.O. *et al.* (2014) Inferring on the intentions of others by hierarchical Bayesian learning. *PLoS Comput. Biol.* 10, e1003810
91. Roth, A.E. and Erev, I. (1995) Learning in extensive-form games - experimental-data and simple dynamic-models in the intermediate-term. *Games Econ. Behav.* 8, 164–212
92. Whiten, A. *et al.* (2009) Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 2417–2428
93. Heyes, C. (2016) Who knows? Metacognitive social learning strategies. *Trends Cogn. Sci.* 20, 204–213
94. Charpentier, C.J. *et al.* (2020) A neuro-computational account of arbitration between choice imitation and goal emulation during human observational learning. *Neuron* 106, 687–699
95. Vélez, N. and Gweon, H. (2021) Learning from other minds: an optimistic critique of reinforcement learning models of social learning. *Curr. Opin. Behav. Sci.* 38, 110–115
96. Wilson, R.C. *et al.* (2014) Humans use directed and random exploration to solve the explore-exploit dilemma. *J. Exp. Psychol. General* 143, 2074–2081
97. Daw, N.D. *et al.* (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879
98. Atlas, L.Y. *et al.* (2016) Instructed knowledge shapes feedback-driven aversive learning in striatum and orbitofrontal cortex, but not the amygdala. *Elife* 5, e15192
99. Boorman, E.D. *et al.* (2013) The behavioral and neural mechanisms underlying the tracking of expertise. *Neuron* 80, 1558–1571
100. Gweon, H. *et al.* (2010) Infants consider both the sample and the sampling process in inductive generalization. *Proc. Natl. Acad. Sci. U. S. A.* 107, 9066–9071
101. Bonawitz, E. *et al.* (2011) The double-edged sword of pedagogy: instruction limits spontaneous exploration and discovery. *Cognition* 120, 322–330
102. Bridgers, S. *et al.* (2020) Young children consider the expected utility of others' learning to decide what to teach. *Nat. Hum. Behav.* 4, 144–152
103. Gweon, H. *et al.* (2014) Sins of omission: children selectively explore when teachers are under-informative. *Cognition* 132, 335–341
104. Kleiman-Weiner, M. *et al.* (2020) Downloading culture.zip: social learning by program induction. In *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*
105. Lyons, D.E. *et al.* (2007) The hidden structure of overimitation. *Proc. Natl. Acad. Sci. U. S. A.* 104, 19751–19756
106. Liu, S. *et al.* (2017) Ten-month-old infants infer the value of goals from the costs of actions. *Science* 358, 1038–1041
107. Shafto, P. *et al.* (2012) Learning from others: the consequences of psychological reasoning for human learning. *Perspect. Psychol. Sci.* 7, 341–351
108. Palminteri, S. *et al.* (2017) The importance of falsification in computational cognitive modeling. *Trends Cogn. Sci.* 21, 425–433
109. Wilson, R.C. and Collins, A.G.E. (2019) Ten simple rules for the computational modeling of behavioral data. *Elife* 8, e49547
110. Mathys, C. *et al.* (2011) A Bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* 5, 39
111. Nassar, M.R. *et al.* (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J. Neurosci.* 30, 12366–12378