Representational exchange in human social learning: Balancing efficiency and flexibility

Charley M. Wu[1](charley.wu@uni-tuebingen.de), Natalia Vélez[2] & Fiery A. Cushman[2]
[1] Human and Machine Cognition Lab, University of Tübingen
[2] Department of Psychology, Harvard University

Abstract

What makes human social learning so powerful? While past accounts have sometimes prioritized finding the single capacity that makes the largest difference, our social learning abilities span a wide spectrum of capacities—from the high-fidelity imitation of behaviors to inferring and learning from hidden mental states. Here, we propose that the power of human social learning lies not within a single capacity, but in our ability to flexibly arbitrate between different computations and to integrate their outputs. In particular, learners can directly copy the demonstrator's actions in the absence of causal insight (*policy imitation*), infer their instrumental values (*value inference*), or infer their model of the world and intrinsic rewards (*belief inference* and *reward inference*). Each of these strategies trades off the cost of computation against the flexibility and compositionality of its outputs. Crucially, we have the capacity to arbitrate and exchange information between these representational formats. Human social learning, we suggest, is powerful not just because of the way it moves information between minds, but also because of the way it flexibly moves information within them.

Introduction

Many animals can learn from each other, but not like us. If chimpanzee social learning were a simple tune carried by a single voice, ours would be the exuberant chorus of a 12-piece ragtime marching band. Naturally, anybody who wants to understand human learning and behavior must confront a central question: Why do they sing, while we *swing*?

Our social learning abilities span a wide spectrum of different capacities, and so researchers have prioritized figuring out which ones matter the most, identifying the "small difference that [...] made a big difference" (Tomasello, Carpenter, Call, Behne, & Moll, 2005). To some, human social learning is powerful primarily because we are uniquely disposed towards high-fidelity imitation of socially observed behaviors, granting us the capability to transmit and innovate upon cultural knowledge (Henrich, 2017; Tennie, Call, & Tomasello, 2009; Boyd & Richerson, 1988). When learning to bake a loaf of bread, for instance, we might imitate each individual action of a master baker, faithfully replicating the same motor responses. We might do this even if we cannot understand the rationale behind her movements and choices (Lyons, Young, & Keil, 2007); Indeed, the master herself might be unaware of the rationale, having inherited some techniques through cultural transmission (Derex, Bonnefon, Boyd, & Mesoudi, 2019; Henrich, 2017). In sum, perhaps the key difference is that humans can transmit cultural knowledge of specific behaviors and action representations with precision.

To others, human social learning is powerful because we can copy not just actions, but also more abstract goals, beliefs, and values, which can be re-assembled productively into new behaviors. Unlike actions, these mental states are not directly observable. Thus, copying them depends on our ability to draw rich social inferences about the unobservable mental states of other people (Jara-Ettinger, 2019; Apperly, 2010; Gweon, 2021; Scott-Phillips, 2017; Strachan, Curioni, Constable, Knoblich, & Charbonneau, 2020). On this view, the uniqueness of human social learning is owed less to high fidelity copying and transmission of concrete behaviors, but rather, shifts the focus to our ability to reconstruct abstract knowledge and values (Sperber, 2006; Morin, 2016). When observing someone bake a loaf of bread, we can acquire knowledge about the leavening power of yeast and extensibility of well-developed gluten, and can emulate goals, such as achieving airy expansion of the loaf in the oven. In sum, perhaps the key difference is that humans can both infer and emulate the hidden rationale of others' actions.

These are both plausible candidates. Compared to other primates, humans are more disposed to high-fidelity imitation (Horner & Whiten, 2005; McGuigan, Whiten, Flynn, & Horner, 2007) and have more sophisticated capacities for mental-state inference (Tomasello et al., 2005; Herrmann, Call, Hernàndez-Lloreda, Hare, & Tomasello, 2007). Nobody seriously believes that human social learning is limited exclusively to one form, whether copying actions, or learning abstract, generative structure. The debate is over which form predominates: Both instruments are in the band, but which carries the melody?

At its heart, our proposal rejects the premise of this question. We propose that the power of human social learning comes not from any single instrument, but from their harmonization. We don't mean this in the banal sense that you can blow harder on two horns than one. Rather, like a musical arrangement, the things we learn must be effectively and efficiently combined. Humans are, then, uniquely masterful composers—social learners with an unrivaled ability to integrate across different levels of representation, from specific behaviors to the hidden generative structure behind them.

Our argument is structured around an analogy to *non*-social learning. It is an apt analogy because here too, humans have a variety of instruments at our disposal. (In fact, we shall argue, the representations involved in social and non-social learning are the same.) Sometimes we rely on habit, recycling past actions, or repeating what has been rewarded in the past. Other times, we plan new actions based on a representation of their likely future consequences. Decades of psychological theorizing were spent arguing over which of these things humans do, or which is more important (e.g., Tolman, 1948; Skinner, 1950). But there is a growing recognition that the essence of human intelligence is not contained within one of these methods individually, but rather in the way that we arbitrate and integrate between them (Huys et al., 2015; Russek, Momennejad, Botvinick, Gershman, & Daw, 2017; Kool, Cushman, & Gershman, 2018; Keramati, Smittenaar, Dolan, & Dayan, 2016; Cushman & Morris, 2015). A baker, for instance, relies at times on skills "in the hands"—a certain way of shaping a loaf, which is a learned behavior that has solidified into habit through countless hours of practice. At other times, she relies on domain knowledge and goals that allow her to adapt to variations in the ingredients, humidity, temperature, and so forth—or to plan out an entirely new recipe. A debate over whether "habit" or "planning" is more important misses the point: What is most remarkable is her ability to compose these elements into a whole that is both practiced and productive; both efficient and flexible (Botvinick & Weinstein, 2014; Rozenblit & Keil, 2002).

We argue that the same is true for social learning. High-fidelity imitation is a close homolog to certain forms of cheap-but-inflexible learning and decision-making, such as habit. Meanwhile, mentalizing and emulation are a close homolog to goal-directed planning. The first part of this paper describes these homologies in detail. Just as a baker prepares her loaf in a way that integrates across diverse processes of decision-making, the baker's apprentice faces the task of learning representations at multiple levels, from concrete actions to abstract knowledge and goals. To be successful, she must arbitrate between imitation and emulation for each part of the bread-making process, and integrate these learned representations with her own pre-existing knowledge, skills and goals, across multiple levels of representation. The heart of our paper addresses the processes of arbitration and integration, which is essential to the power of human social learning.

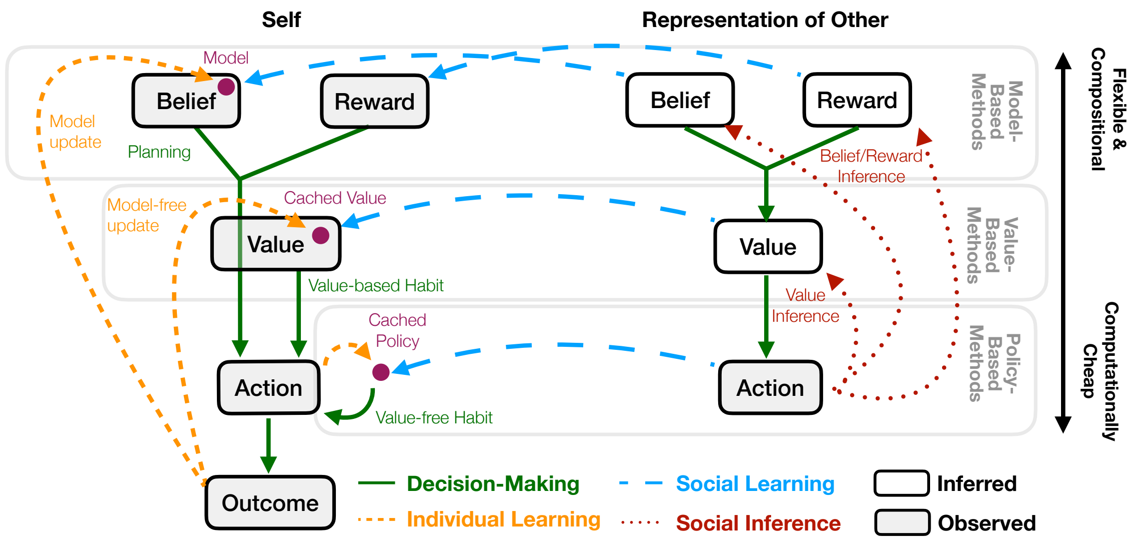

Figure 1. Different forms of social and individual learning share a similar trade-off involving computational complexity and flexibility, which both increase as we ascend the decision-making hierarchy (green solid arrows). *Left*: In individual learning (orange dashed arrows), we can update a cached policy, a cached value representation, or our model of the environment. *Right*: Similarly, different forms of social learning (blue long dashed arrows) draw upon different representational formats: Directly adopting the actions or policy of another agent, adopting their inferred value representations, or adopting their inferred beliefs or intrinsic reward function. Value-based and model-based methods of social learning require inferring hidden mental states from observed actions (red dotted arrows), which incur added computational costs. However, they also afford increased flexibility, since learning from these primitives of the decision-making process allows us to deploy them in an adaptive and compositional fashion.

For simplicity, we focus on observational learning, where we perceive a person's action and the setting in which they are performing it, but without direct access to the underlying causes of their action (e.g., via explicit communication). We will also assume a relatively naïve observer is observing a relatively experienced expert, and attempts to learn how to act in more expert ways themselves (like a baker and her apprentice). Clearly human social learning involves much more than this specific form of observational learning. It also involves teaching, talking, rewarding and punishing, and much more. Nevertheless, our case study of observational learning aims to inform theories of human social learning more broadly.

Mechanisms of social learning and decision-making

There is a close relationship between different forms of individual and social learning (Najar, Bonnet, Bahrami, & Palminteri, 2020; Morris & Cushman, 2018). Theories of observational social learning have been broadly divided into two approaches: Imitation and emulation. With imitation, the learner copies the observed *action* (N. E. Miller & Dollard, 1941; Bandura, 1962; Heyes, 2001).

With emulation, the learner decomposes the observed behavior into a set of reconstructed primitives. This could be the *value* of the action, or it could be the *beliefs* and ultimate *rewards* that dictate the expected value of an action with respect to some goal (Tomasello, Davis-Dasilva, Camak, & Bard, 1987; Whiten & Ham, 1992). This defines a hierarchy of inferences an observer can draw from a demonstrator's actions, and a corresponding set of mental representations that the observer can potentially adopt via social learning.

Analogously, when deciding how to act, a person can rely on any of several strategies. They can draw directly from a representation of which *action* to perform (i.e., a cached policy, or stimulus-response repertoire). They can draw from a representation of the instrumental *value* of the actions available to them (from cached values). Or, they can select potential actions by considering the likely future consequences of those actions given their *beliefs* about the environment and intrinsic *rewards* (i.e., model-based planning). This defines a hierarchy of individual decision-making mechanisms[1].

In this section, we will introduce a taxonomy of social learning that draws an analogy between well-understood mechanisms of individual learning (Fig. 1, left), and the hierarchy of inferences that an observer can draw in social learning (Fig. 1, right). When an individual performs an action, observes a change in the world, and receives some reinforcement, she faces at least three ways to learn from this. She can update her *cached policy*, her representation of the *value* of the action performed, or her *world model*. Similarly, when an individual observes someone else's actions, there are several ways that she could use that observation to guide her own behavior. At the simplest level, she could take on that person's action as her own, directly imitating the person's behavioral policy (*policy imitation*). Moving up the hierarchy, she could infer the unobservable mental states that produced the action—namely, she could use her inferences about that person's value to update her own representations of the instrumental value of certain actions (*value inference*), her inferences about that person's beliefs to update her own beliefs about the world (*belief inference*), or even her inferences about what the person finds intrinsically rewarding to update her own experience of reward (Cushman & Morris, 2015; Zaki, Schirmer, & Mitchell, 2011).

This connection is useful because it allows us to export well-studied features of individual learning to better characterize social learning. We focus on three key ideas. First, these mechanisms of individual learning exhibit different tradeoffs between computational efficiency, on one hand, and both flexibility and compositionality, on the other. Relying on cached representations of policy or value is computationally cheap, but it is less flexible and less compositional than full model-based planning, which can adapt to new goals and contexts, and can reassemble representational elements into novel behavior. Second, due to this tradeoff, people arbitrate between different forms of learning and decision-making, selecting computationally cheap methods when possible and more complex approaches when necessary. Third, complex real world behaviors depend on the adaptive integration of different representations (e.g., policies, values, and beliefs). This in turn depends on processes like planning and inference, which can transfer information across levels of representation. In the remainder of the chapter we address the implications of each of these points for theories of social learning.

Imitation and policy caching

In the simplest case, observing a behavior by another individual will often make us more likely to also adopt it (Fig. 2). In other words, social learning often involves action imitation (Hoppitt &

[1]Following convention in reinforcement learning (Sutton & Barto, 2018), we use "reward" for intrinsic objectives, such as eating food, and "value" for things of learned instrumental utility, such as planting seeds.

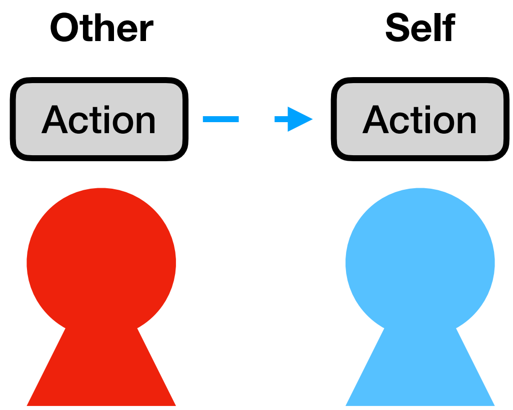

Figure 2. Schematic of policy imitation

Laland, 2013; Whiten & Ham, 1992; Legare & Nielsen, 2015). You might simply choose a restaurant based on its popularity (Boyd & Richerson, 1988), or you might follow one of many widely studied imitation biases to selectively imitate the actions of prestigious (Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019) or previously successful individuals (Atkisson, O'Brien, & Mesoudi, 2012; Rendell et al., 2010). Paired with an appropriate rate of even random variability (akin to "mutations") in the copying process, imitation has been proposed to be a key driver of human cultural evolution (Heyes, 2018; Tennie et al., 2009; Henrich, 2017).

Imitation is often considered computationally cheaper and simpler compared to other forms of social learning, since no inference is needed about the other person's goals or intentions (Catmur, Walsh, & Heyes, 2009; Heyes, 2002). Rather, their behavioral policy can simply be copied verbatim. Thus, a useful analogy can be made to individual learning, where we often deploy a "cached policy" in the form of habitual behaviors (Gershman, 2020; K. J. Miller, Shenhav, & Ludvig, 2019; Dezfouli & Balleine, 2013). Habits are patterns of behaviors that are learned through exercise and repetition, rather than reinforced through rewarding outcomes (Thorndike, 1932). For instance, we can take our usual route to work in the morning or order our usual meal at our favorite restaurant, without having to iterate over possible plans or weighing the benefits of different options on the basis of expected value. Sometimes referred to as "amortization" (Dasgupta, Schulz, Goodman, & Gershman, 2018), we can simply avoid costly computations by caching and redeploying solutions that have worked in the past.

In exchange for its simplicity, pure imitation lacks flexibility. Suppose you meet a friend for lunch at a restaurant with a confusing and sprawling menu. Since your friend is a regular, you could always copy their lunch order, saving you the trouble of blundering your way through trial and error. However, you are likely to reach the limits of this strategy on subsequent visits, where you may miss out on new, higher-reward options (e.g., a changing daily special), or you may have no idea what to get when your friend's usual order is out of stock. Thus, compared to more sophisticated forms of social learning, imitation generalizes poorly and lacks compositionality, since one cannot pluck an action out of its original context and expect it to be equally useful when placed in a new situation.

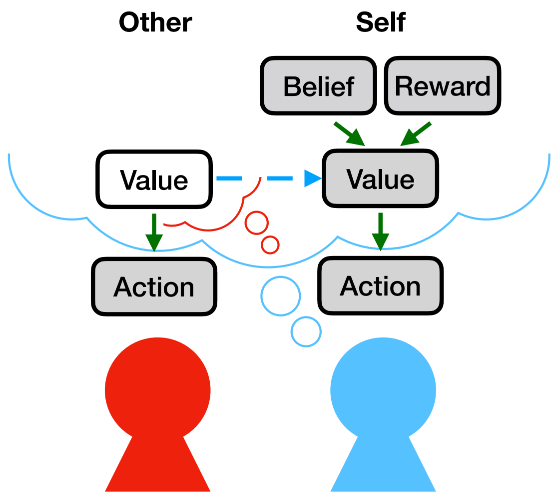

Figure 3. Schematic of value inference

Value Inference and Cached Value

Another form of observational learning is to infer the instrumental value another person assigns to different actions by observing their behavior, and then to adopt this value representation oneself (Fig. 3; Ho, MacGlashan, Littman, & Cushman, 2017; Collette, Pauli, Bossaerts, & O'Doherty, 2017; Jern, Lucas, & Kemp, 2017). Inferring values from social observation is closely related to the well-established use of cached value representations or value-based habits in individual learning (Daw, Niv, & Dayan, 2005; Keramati et al., 2016; Kool, Cushman, & Gershman, 2018; Solway & Botvinick, 2012). Like the use of a cached policy, caching value representations allows an agent to re-use costly computations, although this process takes place one step higher up the decision-making hierarchy (Fig. 1), since observed actions first need to be "unpacked" into inferred value representations. Developing a representation of instrumental value—such as the instrumental value of trading a knight for a bishop in chess—is generally costly since it requires some amount of thinking, learning, and practice. Having earned this insight, it makes sense to "cache" it for reuse in future games. While not as cheap as caching policies, caching value representations offers superior generalizability and flexibility.

We can formally characterize social value inference as inverse reinforcement learning (IRL Jara-Ettinger, 2019). Standard reinforcement learning (RL) models (Sutton & Barto, 2018) are used to understand how agents (biological or artificial) learn through interactions with the environment in non-social settings. As the agent interacts with the environment, they receive rewards, which update the value representations for different states and actions, which in turn guides future behavior. IRL inverts this model using Bayes' rule to infer hidden value representations given observed actions.

While IRL illuminates how humans reason about others' preferences, goals, and beliefs (Collette et al., 2017; Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017; Jern et al., 2017), it is computationally costly. For most interesting problems IRL is computationally intractable (Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016; Vélez & Gweon, 2021). Nevertheless, as a rational framework it can be used to uncover inductive biases that simplify the required computations. For instance, the *principle* *of efficient action* (Gergely & Csibra, 2003; Jara-Ettinger, Gweon, Tenenbaum, & Schulz, 2015; Liu, Brooks, & Spelke, 2019) assumes that other agents are acting in the most efficient manner towards achieving their aims, greatly constraining the hypothesis space for IRL inference.

A key advantage of cached value, rather than cached policy, is it's superior generalization. Consider, for instance, an experiment in which participants are given repeated choices between two options that pay off with (A) 20% and (B) 30% probability, and also between options that pay off with (C) 70% and (D) 80% probability . Naturally, the proper policy is to choose B and D, respectively. But now suppose that a person is given the novel choice between B and C. Having cached a *policy* representation, they might choose B over C—after all, they have frequently chosen B, and rarely C (K. J. Miller et al., 2019; Hayden & Niv, 2021). In contrast, having cached a *value* representation, they should choose C over B, because C has been associated with greater historical rewards. This captures the key respect in which value representation affords greater compositionality than policy representation.

The same point applies to social learning. When imitating behavior at the level of policy representations, one does not distinguish between a person who chooses to eat a *dax* (some unknown fruit) when they could have eaten a ripe peach, versus someone who chooses to eat a *dax* given the alternative of a dirty hotdog. When learning from inferred value representations, however, one will likely assign greater value to the *dax* in the first case than the second. This, in turn, will help the learner to make adaptive choices about when to eat a *dax* herself—a key form of flexibility.

Importantly, cached value can be assigned not only to actions, but also to more abstract representations such as goal-subgoal relations (Cushman & Morris, 2015; Keramati et al., 2016; Maisto, Friston, & Pezzulo, 2019; Botvinick & Weinstein, 2014). For instance, the superordinate goal "make coffee in the morning" can be assigned value, but can also induce a set of value representations over subordinate goals: When valuing morning coffee, one should assign value to grinding beans, heating water, and so forth. This kind of cognitive architecture is naturally suited to "goal emulation" (Tomasello, 1996), in which the observer imputes goals to the actor and then adopts those goals, but uses novel planning to derive her own policy for attaining the goal.

Belief Inference and Model-based Planning

Another method of social learning is to update one's own model of the world (i.e., beliefs) when observing another person's actions (Fig. 4). For instance, you might infer that it is likely to rain later by observing someone on the street carrying an umbrella. This observation could then inform you that you can skip watering the garden in the evening or that you should reschedule a planned picnic in the park.

Belief inference affords great flexibility and compositionality. Copying another person's actions, or adopting their representations of instrumental value, may not always be useful to you: those actions and values depend on the actor's unique circumstances, goals, or desires. In contrast, facts about the world are equally true for all of us. This affords maximal flexibility. Socially learned beliefs can be seamlessly composed with existing beliefs, and via planning, used to generate novel actions that reflect the specific circumstances, abilities, and desires of the actor.

Just as with values, an actor's unobserved beliefs can be inferred by Bayes' rule, given observations of their actions (Baker et al., 2017; Jara-Ettinger et al., 2016; Shafto, Goodman, & Frank, 2012). However, this may be computationally costlier than inferring the more direct linkage between actions and instrumental values. It is often possible to infer a person's values from their actions without considering their beliefs; for instance, concluding that they like apples when you see them bite into

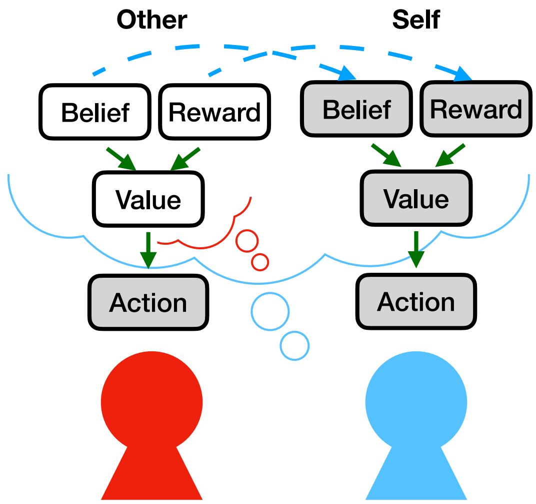

Figure 4. Schematic of belief and reward inference

one. But it is much harder to infer a person's beliefs from their action without performing a joint inference over their values (or rewards) as well. If a person opens the cabinet, you cannot know what they *believe* is in the cabinet without a hypothesis about what they *want* to retrieve from it. This may help to explain why very young infants are relatively adept at inferring the goals or values of other people, whereas robust inference about the beliefs of others comes much later (Gergely & Csibra, 2003).

There is a natural relationship between model-based social inference and model-based planning in individual learning, both in terms of the underlying representations and the computational demands of using them (K. J. Miller, Botvinick, & Brody, 2017; Vikbladh et al., 2019). People often make decisions by computing the expected values of candidate actions given a model of their likely outcomes—in other words, their world model or beliefs. This is more computationally demanding than working from cached value or cached policy, but it affords the greatest degree of behavioral flexibility. Whereas cached policies and cached values are insensitive (without update) to changes in one's knowledge about the world or one's preferences, the process of planning over a world model maintains full sensitivity.

Reward Inference

When inferring the mental causes of observed actions, we can impute not only people's instrumental values ("she thinks it is useful to buy broccoli at the grocery store"), but also their intrinsic reward specification ("she loves the taste of broccoli"). Intrinsic rewards have been argued to be an important and useful target for social learning (Ho et al., 2017), since they may reveal elements of social partners' reward specifications that confer valuable fitness benefits (e.g., adopting a colleague's healthier food preferences). And, consistent with this possibility, a large social psychological literature documents examples of adopting others' preferences as "normative conformity" or "internalization" (Morris & Cushman, 2018). For instance, learning which faces other people find most attractive will influence one's own preferences (Zaki et al., 2011). Like the inference of beliefs and instrumental values, the inference of intrinsic rewards can be accomplished by IRL.

What are the costs and benefits of socially learning intrinsic rewards (e.g., "broccoli is delicious")? There are two sources of computational cost: inference and decision-making. Like inferring instrumental values or beliefs, inferring rewards can be accomplished by IRL, which requires costly computation. Then, after adopting a new reward, this can influence behavior in a variety of ways. The agent can use computationally expensive model-based methods to immediately update their policy via planning, or they can use computationally cheap model-free methods to gradually update their policy. Current research indicates that people can also use methods that blend model-free and model-based elements (Momennejad et al., 2017). Meanwhile, there are key benefits of socially learning new rewards. Like learning new beliefs, this form of learning is maximally compositional in the sense that a new optimal policy can be computed or learned following any arbitrary change to the reward function or the environment.

Interim Summary

Compared to information gained through first hand interaction with the environment, social information has a rich structure built on hidden mental states. Observing your friend order a sandwich, you can infer information about both the world (e.g., the available menu options) and your friend (her values and beliefs). This information is implicitly "packed" into every socially observed action. We have defined a hierarchy of social learning mechanisms that differ in how the learner "unpacks" observed actions to guide her own behavior. Each of these social learning mechanisms trades off the cost of performing the computation against the flexibility and compositionality of its outputs. It is computationally cheap to copy the "packed" information in the same format we observed it—that is, to use our observations of others' actions to guide our own actions directly. But, by "unpacking" the latent structure that gave rise to the action, this information can be productively composed with unique features of our own circumstances (e.g., our goals, abilities, and situational context). For example, you might infer that your friend's choice of sandwich is tastier than other options currently on the menu, but pass it up when the restaurant announces an exciting daily special. Drawing on this framework, our central proposal is that human social learning is powerful not because of any of these mechanisms, but rather because of how information is shared and composed across different mechanisms and levels of representation. Next, we consider how this is accomplished.

Representational exchange in social learning

Our representational exchange framework of social learning (Fig. 1) organizes different forms of social and non-social learning based on the representations involved. This focus on representational levels allows us to address two important problems within the same framework: How humans arbitrate among social learning mechanisms, and how information is exchanged and integrated across representations. Together, arbitration and integration enable humans to produce complex behaviors that efficiently blend cached representations with those that are computed online.

Arbitration

To some, the idea that humans arbitrate among social learning mechanisms may seem puzzling. Humans can make incredible inductive leaps from sparse social information by reasoning about the beliefs and values of other people (Vélez & Gweon, 2021). Why not always use the most sophisticated instrument in our repertoire?

One promising answer to this question comes from resource-rational theories of cognition (Bhui, Lai, & Gershman, 2021; Gershman, Horvitz, & Tenenbaum, 2015; Lieder & Griffiths, 2020). Human minds are fundamentally limited—we have access to limited data, both from the world and from other people, and we have limited time and resources for performing computations. Resource-rational theories formalize how agents should make optimal use of limited minds. A common thread that runs through these theories is that people select strategies by balancing the effort required to perform a computation against the payoffs or precision of its outputs. This general principle shapes the arbitration between model-free habits (cached policy or value) versus model-based planning in individual decision-making (Kool, Gershman, & Cushman, 2018, 2017). In tasks where model-based control generates little advantage in expected reward, people favor computationally, cheaper model-free methods instead. When model-based control reliably produces a reward advantage, the degree to which they rely on it scales with the magnitude of the extra rewards.

We speculate that a similar principle might govern arbitration among social learning mechanisms: One should trade off the cost of performing a computation against the flexibility and compositionality of its outputs. A prediction that follows from this framework is that learners may deploy different social learning mechanisms based on the incentives of the task. Indeed, work on non-social learning and decision-making suggests that people engage in more effortful, controlled forms of reasoning when they are offered higher payoffs (Kool et al., 2017); conversely, when people make decisions under stress (Otto, Raio, Chiang, Phelps, & Daw, 2013), strict time limits (Wu, Schulz, Gerbaulet, Pleskac, & Speekenbrink, 2019), or increased cognitive load (Cogliati Dezza, Cleeremans, & Alexander, 2019), they fall back on cheaper, habitual behaviors instead (Otto, Gershman, Markman, & Daw, 2013; Kool, Cushman, & Gershman, 2018; see Shenhav, Botvinick, & Cohen, 2013 for a review). Similarly, we might expect people to engage in more costly forms of social learning (e.g., belief inference) when they are incentivized to get it right, and to fall back on less costly forms (e.g., policy imitation) under increased task demands. If your friend orders a sandwich, you *could* compare the sandwich to other menu options to infer their preferences and beliefs in order to flexibly plan the best possible order—but if you are tired and need a quick bite to eat, you might get whatever she's having instead.

Further, more sophisticated inferences may miss the mark when they are based on insufficient evidence or when our model of the world is incomplete (Gigerenzer & Gaissmaier, 2011). Policy imitation does not only save computational costs; it also provides a way for learners to hedge their bets. For instance, if your friend demonstrates how to operate a very complex, delicate, and expensive Magnetic Resonance Imaging (MRI) scanner, it may be safer to faithfully reproduce her button presses than to use your limited knowledge to infer which steps can be skipped. Recent behavioral and neuroimaging evidence suggests that when comparing choice imitation against value inference (called goal emulation in the paper), people preferred the strategy that produced the most reliable predictions about outcomes (Charpentier, Iigaya, & O'Doherty, 2020). Activity in the bilateral temporoparietal junction and right ventrolateral prefrontal cortex tracked the trial-by-trial reliability of predictions generated through value inference; this activity is hypothesized to provide a control signal, regulating which of these predictions guides behavior.

In sum, humans face two fundamental challenges when arbitrating between social learning mechanisms: whether to engage in costlier forms of inference at all, and which outputs to use to guide behavior. How humans navigate these challenges has been largely unexplored (but see Charpentier et al., 2020). We hypothesize that humans decide whether to deploy costlier forms of inference by balancing the cost of the computation against the precision or payoff of its outputs, and they adjudicate the outputs of different computations based on their reliability.

Integration

So far we have emphasised three "horizontal pathways" for social learning (Fig. 1): The direct copying of policy, the adoption of imputed values, and the adoption of imputed beliefs and rewards. Having discussed how information is transferred between minds, we now turn to the "vertical pathways" for exchanging information across representational formats within one's mind (Cushman, 2020). This includes ascending from observed actions to imputed mental states via inference, and also descending from mental states into actions via value-guided decision-making. The success of human social learning is not only in our ability to acquire the right form of social information, but also to productively exchange information between levels of representation and to integrate it with our own experiences.

Descending: Amortization and offline planning

The simplest representational exchange is ordinary decision-making. For instance, planning involves the compression of one's model of the environment into a value representation, and then implementing a policy to select actions on the basis of value (Sutton & Barto, 2018). Each of these steps creates new representations that can be stored or "cached" for reuse in similar future episodes (Cushman & Morris, 2015; Keramati et al., 2016; Maisto et al., 2019), such that the computational burden can be distributed over time (i.e., amortization; Dasgupta et al., 2018).

This process can occur without actually taking any actions at all, but merely through the act of planning and simulating action (Gershman, Markman, & Otto, 2014). For instance, a downhill skier can form a mental model of the course and visualize the process of navigating it in the safety of her hotel room, acquiring cached representations through mental simulations. There is evidence that this kind of "offline planning" can occur without conscious effort, and even during sleep (Momennejad, Otto, Daw, & Norman, 2018; Foster, 2017), helping to generate compressed representations that facilitate rapid performance on the course.

These ideas naturally extend to social learning, since a skier can also learn by observing an expert performance. If we adopted the view that "imitation" and "emulation" are exclusively competing strategies, then our skier would face a difficult choice. Should she copy the precise actions of an expert, hoping that the differences between their bodies, skills, and mountains are not too great? Or, should she fully decompose the observed behavior into its causal primitives—the beliefs, rewards, and values that shaped the expert's performance? This offers more flexibility, but it's difficult to imagine she could make it down the mountain at any reasonable speed while computing the best moves from first principles. Offline planning may, then, play a key role in making emulation feasible. Rather than having to recompute her plan while on the slope, she can integrate whatever she has learned from the expert with her own knowledge in advance, through offline planning.

Of course, we don't always "unpack" observed behavior into its most elemental components and then fully re-plan new policies. Rather, we likely employ a mixture of different social strategies, guided by the principles of arbitration: using cheap methods whenever possible and more costly methods when more reliable. Offline planning—like ordinary experience—-can be used to re-assemble these components into a new and better trajectory down the hill.

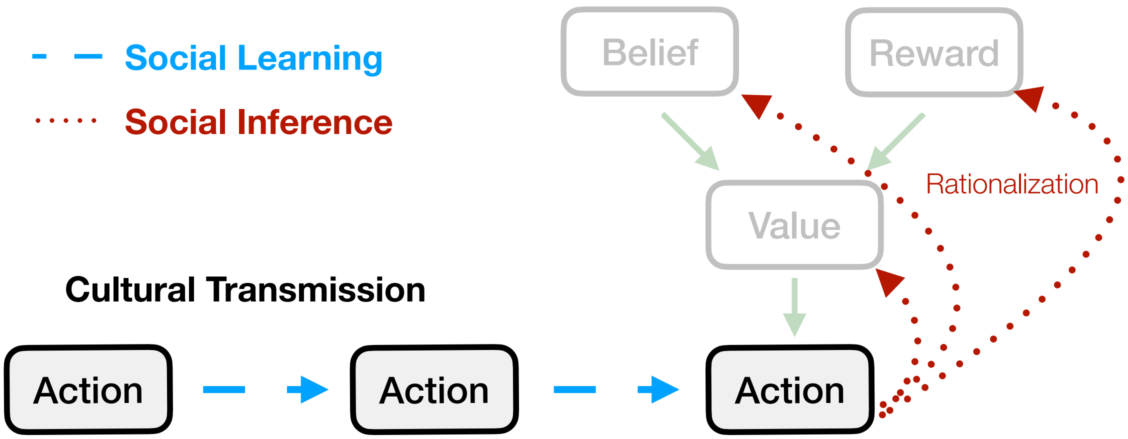

Figure 5. "Rationalization" of cultural knowledge. Behaviors may be transmitted through cultural norms, but because observed actions are rich in implicit information, they can be "unpacked" into imputed mental states. These mental states may not belong to the other individual (who directly acquired the behavior through imitation), yet they provide a useful fiction, exchanging an inflexible cultural norm for flexible and compositional representations to solve new problems.

Ascending: Inference and rationalization

When we observe another person's actions, we can infer the hidden mental states that gave rise to those actions, including their beliefs, reward function, and computations of instrumental value. This is a basic form of "ascent" in our model (Fig. 1). And, in the social cognition literature, this kind of thing is said so frequently that it can be easy to overlook it's hidden premise: That, when we observe another person's actions, those actions really were caused by their beliefs, reward function, and computations of instrument value—that is, by rational planning.

But, often, they aren't. People's actions are also the product of habit (cached values or cached policies), instinct (innate behavioral responses), or conformity to cultural norms (when they imitate what other people do without reasoning about why they do it). In these cases, what would it mean if an observer imputed beliefs, a reward function, and computations of instrumental value to the actor? It would mean that the observer is constructing a fiction—a "rationalization" of behavior—which can nevertheless be quite useful (Cushman, 2020). This fiction furnishes a key method of representational exchange, extracting implicit information from the cached policies or values of other people.

Suppose, for instance, that an aspiring baker wishes to improve the flavor of her loaves. She notices that in her culture people let their dough rise overnight, and she imputes the belief that this is a superior method—perhaps because the cool temperatures allow flavor to build during a longer fermentation. So, she tries putting her loaves in especially cool spots, even during the daytime. Now, it might be the case that in her culture, nobody knows why they let the bread rise overnight, they just do it because "that's how it's done". Nevertheless, by imputing beliefs, values and rational choice, the aspiring baker might learn useful information from cultural practices that she can generalize productively.

The rationalization of cultural practices is ubiquitous. Its basic structure is depicted in Figure 5: A cultural norm is transmitted at the level of cached policy, but then later it is rationalized, yielding putative representations of values and beliefs. Parents and schoolteachers constantly find themselves attempting to "make sense" out of cultural practices for inquisitive children. Sometimes we are honest with ourselves, acknowledging that nobody really knows why we do things a certain way, although there may be implicit wisdom we can extract nevertheless. But just as often, we are less honest with ourselves (and with children), acting as if we had either chosen this behavior—or some forgotten designer had created it—for precisely the reasons we articulate to the child. In this case we are constructing a fiction, but a potentially useful one. The resulting representations may enable flexible and compositional thinking where, formerly, an inflexible cultural norm prevailed.

Of course, we rationalize not only other's behaviors, and "culture" more generally, but also our own behavior. Insofar as our own behavior is caused by non-rational processes (e.g., habits, instincts or norms), we can falsely impute reasoning processes to ourselves, assigning beliefs, values, and rewards as if the behaviors we perform were rational. Insofar as our behaviors are often guided by social learning, this is another pathway by which socially acquired information can propagate across levels of representation within our own mind (Cushman, 2019). For instance, having learned the motor routines involved in bread-making (a form of policy imitation), we can rationalize our own actions and thus extract useful values or beliefs about the bread-making process for further innovation.

Summary: Balancing productivity and reuse

Complex human behavior strikes a remarkable balance between flexible productivity and efficient reuse. This is true of the way we learn theories (Rozenblit & Keil, 2002), make plans (Cushman & Morris, 2015; Keramati et al., 2016; Solway et al., 2014), and speak (O'Donnell, 2015). This balance of productivity and reuse can be viewed as a variety of resource-rational cognition (Lieder & Griffiths, 2020). If we brought all of our knowledge and thinking to bear on every single problem, we'd never be able to make any decisions in time. Thus, wherever possible, we reuse the outputs of computations from relevant past episodes, while devoting our most powerful but precious cognitive resources to the specific elements of a behavior that require revision or improvement (Dasgupta & Gershman, 2021). To develop a complete theory of human social learning, we must understand how humans decompose observed behavior into disparate elements, each of which is represented at the appropriate level for efficient composition.

Conclusion

Social learning is central to the human experience. But rather than trying to pick out a single distinguishing feature of what makes humans special, we have argued the distinction is in how the different mechanisms of social learning interact. In this we have drawn inspiration from the literature on non-social decision-making, where there is emerging consensus that human intelligence arises from the productive cooperation of diverse strategies for learning and decision-making (Kool, Cushman, & Gershman, 2018; Huys et al., 2015; Solway & Botvinick, 2015). This is a natural source of inspiration because the hierarchy of inferences that we can draw from during observational social learning mirrors the hierarchy of learning and decision-making systems available to an individual. Both of these hierarchies can be characterized as points on a tradeoff between computational efficiency on the one hand and representational flexibility and compositionality on the other. The connection between social learning and individual decision-making is also natural because we integrate what we have learned with our pre-existing policies, values and beliefs. To do this, we exchange information between different formats of representation. This involves “descending” pathways (moving from more flexible and compositional representations towards compressed policy-relevant representations), as well as “ascending” pathways (extracting information implicit in policy-relevant representations into more flexible and compositional elements).

We study social learning because we want to understand real human behaviors—the kinds we perform every day, such as baking a loaf of bread or choosing what to eat. These behaviors are paradigmatic of human intelligence, partly because we are able to blend strategies we've learned from others with strategies we've developed ourselves. Yet they are also paradigmatic because they involve skills and representations spanning from very specific, concrete behaviors to very abstract, general principles. Our ability to build harmony across these levels is essential to the virtuosic performances of the human mind.

Acknowledgements

CMW is supported by the German Federal Ministry of Education and Research (BMBF): Tübingen AI Center, FKZ: 01IS18039A and funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy–EXC2064/1–390727645. NV is supported by grant 8K00MH125856-02 from the National Institute of Mental Health. FAC is supported by grant N00014-19-1-2025 from the Office of Naval Research and grant 61061 from the John Templeton Foundation.

References

Apperly, I. (2010). *Mindreaders: The cognitive basis of "theory of mind"*. Psychology Press.

Atkisson, C., O'Brien, M. J., & Mesoudi, A. (2012). Adult learners in a novel environment use prestige-biased social learning. *Evol. Psychol.*, *10*(3), 519–537.

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 1–10.

Bandura, A. (1962). Social learning through imitation. *Nebr. Symp. Motiv.*, *330*, 211–274.

Bhui, R., Lai, L., & Gershman, S. J. (2021). Resource-rational decision making. *Current Opinion in Behavioral Sciences*, *41*, 15–21.

Botvinick, M., & Weinstein, A. (2014). Model-based hierarchical reinforcement learning and human action control. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, *369*(1655).

Boyd, R., & Richerson, P. J. (1988). *Culture and the evolutionary process.* University of Chicago Press.

Catmur, C., Walsh, V., & Heyes, C. (2009). Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, *364*(1528), 2369–2380.

Charpentier, C. J., Iigaya, K., & O'Doherty, J. P. (2020). A neuro-computational account of arbitration between choice imitation and goal emulation during human observational learning. *Neuron*, *106*(4), 687–699.e7.

Cogliati Dezza, I., Cleeremans, A., & Alexander, W. (2019). Should we control? the interplay between cognitive control and information integration in the resolution of the exploration-exploitation dilemma. *J. Exp. Psychol. Gen.*, *148*(6), 977–993.

Collette, S., Pauli, W. M., Bossaerts, P., & O'Doherty, J. (2017). Neural computations underlying inverse reinforcement learning in the human brain. *Elife*, *6*.

Cushman, F. (2020). Rationalization is rational. *Behav. Brain Sci.*, *43*, e28.

Cushman, F., & Morris, A. (2015). Habitual control of goal selection in humans. *Proc. Natl. Acad. Sci. U. S. A.*, *112*(45), 13817–13822.

Dasgupta, I., & Gershman, S. J. (2021). Memory as a computational resource. *Trends in Cognitive Sciences*.

Dasgupta, I., Schulz, E., Goodman, N. D., & Gershman, S. J. (2018). Remembrance of inferences past: Amortization in human hypothesis generation. *Cognition*, *178*, 67–81.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.*, *8*(12), 1704–1711.

Derex, M., Bonnefon, J.-F., Boyd, R., & Mesoudi, A. (2019). Causal understanding is not necessary for the improvement of culturally evolving technology. *Nat Hum Behav*, *3*(5), 446–452.

Dezfouli, A., & Balleine, B. W. (2013). Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized. *PLoS Comput. Biol.*, *9*(12), e1003364.

Foster, D. J. (2017). Replay comes of age. *Annu. Rev. Neurosci.*, *40*, 581–602.

Gergely, G., & Csibra, G. (2003). *Teleological reasoning in infancy: the naıve theory of rational action* (Vol. 7) (No. 7).

Gershman, S. J. (2020). Origin of perseveration in the trade-off between reward and complexity. *Cognition*, *204*, 104394.

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278.

Gershman, S. J., Markman, A. B., & Otto, A. R. (2014). Retrospective revaluation in sequential decision making: a tale of two systems. *J. Exp. Psychol. Gen.*, *143*(1), 182–194.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annu. Rev. Psychol.*, *62*, 451–482.

Gweon, H. (2021). *Inferential social learning: How humans learn from others and help others learn.*

Hayden, B. Y., & Niv, Y. (2021). The case against economic values in the orbitofrontal cortex (or anywhere else in the brain). *Behav. Neurosci.*, *135*(2), 192–201.

Henrich, J. (2017). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter.* Princeton University Press.

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evol. Hum. Behav.*, *22*(3), 165–196.

Herrmann, E., Call, J., Hernàndez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science*, *317*(5843), 1360–1366.

Heyes, C. (2001). Causes and consequences of imitation. *Trends Cogn. Sci.*, *5*(6), 253–261.

Heyes, C. (2002). Transformational and associative theories of imitation. *Imitation in animals and artifacts.*, *607*, 501–523.

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking.* Harvard University Press.

Ho, M. K., MacGlashan, J., Littman, M. L., & Cushman, F. (2017). Social is special: A normative framework for teaching with and learning from evaluative feedback. *Cognition*, *167*, 91–106.

Hoppitt, W., & Laland, K. N. (2013). *Social learning: An introduction to mechanisms, methods, and models.* Princeton University Press.

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (pan troglodytes) and children (homo sapiens). *Anim. Cogn.*, *8*(3), 164–181.

Huys, Q. J. M., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., ... Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proc. Natl. Acad. Sci. U. S. A.*, *112*(10), 3098–3103.

Jara-Ettinger, J. (2019). Theory of mind as inverse reinforcement learning. *Current Opinion in Behavioral Sciences*, *29*, 105–110.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends Cogn. Sci.*, *20*(8), 589–604.

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition*, *140*, 14–23.

Jern, A., Lucas, C. G., & Kemp, C. (2017). People learn other people's preferences through inverse decision-making. *Cognition*, *168*, 46–64.

Jiménez, Á. V., & Mesoudi, A. (2019). Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Communications*, *5*(1), 20.

Keramati, M., Smittenaar, P., Dolan, R. J., & Dayan, P. (2016). Adaptive integration of habits into depth-limited planning defines a habitual-goal-directed spectrum. *Proc. Natl. Acad. Sci. U. S. A.*, *113*(45), 12868–12873.

Kool, W., Cushman, F. A., & Gershman, S. J. (2018). Chapter 7 - competition and cooperation between multiple reinforcement learning systems. In R. Morris, A. Bornstein, & A. Shenhav (Eds.), *Goal-Directed decision making* (pp. 153–178). Academic Press.

Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-Benefit arbitration between multiple Reinforcement-Learning systems. *Psychol. Sci.*, *28*(9), 1321–1333.

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

Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends Cogn. Sci.*, *19*(11), 688–699.

Lieder, F., & Griffiths, T. L. (2020). *Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources* (Vol. 43).

Liu, S., Brooks, N. B., & Spelke, E. S. (2019). Origins of the concepts cause, cost, and goal in prereaching infants. *Proc. Natl. Acad. Sci. U. S. A.*, *116*(36), 17747–17752.

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proc. Natl. Acad. Sci. U. S. A.*, *104*(50), 19751–19756.

Maisto, D., Friston, K., & Pezzulo, G. (2019). Caching mechanisms for habit formation in active inference. *Neurocomputing*, *359*, 298–314.

McGuigan, N., Whiten, A., Flynn, E., & Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3- and 5-year-old children. *Cogn. Dev.*, *22*(3), 353–364.

Miller, K. J., Botvinick, M. M., & Brody, C. D. (2017). Dorsal hippocampus contributes to model-based planning. *Nat. Neurosci.*.

Miller, K. J., Shenhav, A., & Ludvig, E. A. (2019). Habits without values. *Psychol. Rev.*, *126*(2), 292–311.

Miller, N. E., & Dollard, J. (1941). *Social learning and imitation* (Vol. 55). New Haven, CT: Yale University Press.

Momennejad, I., Otto, A. R., Daw, N. D., & Norman, K. A. (2018). Offline replay supports planning in human reinforcement learning. *Elife*, *7*.

Momennejad, I., Russek, E. M., Cheong, J. H., Botvinick, M. M., Daw, N. D., & Gershman, S. J. (2017). The successor representation in human reinforcement learning. *Nat Hum Behav*, *1*(9), 680–692.

Morin, O. (2016). *How traditions live and die.* Oxford University Press.

Morris, A., & Cushman, F. (2018). A common framework for theories of norm compliance. *Soc. Philos. Policy*, *35*(1), 101–127.

Najar, A., Bonnet, E., Bahrami, B., & Palminteri, S. (2020). The actions of others act as a pseudo-reward to drive imitation in the context of social reinforcement learning. *PLoS Biol.*, *18*(12), e3001028.

O'Donnell, T. J. (2015). *Productivity and reuse in language: A theory of linguistic computation and storage.* MIT Press.

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. *Psychol. Sci.*, *24*(5), 751–761.

Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proc. Natl. Acad. Sci. U. S. A.*, *110*(52), 20941–20946.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., … Laland, K. N. (2010). Why copy others? insights from the social learning strategies tournament. *Science*, *328*(5975), 208–213.

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: an illusion of explanatory depth. *Cogn. Sci.*, *26*(5), 521–562.

Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J., & Daw, N. D. (2017). Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS Comput. Biol.*, *13*(9), e1005768.

Scott-Phillips, T. C. (2017). A (simple) experimental demonstration that cultural evolution is not replicative, but reconstructive — and an explanation of why this difference matters. *J. Cogn. Cult.*, *17*(1-2), 1–11.

Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others: The consequences of psychological reasoning for human learning. *Perspect. Psychol. Sci.*, *7*(4), 341–351.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240.

Skinner, B. F. (1950). Are theories of learning necessary? *Psychol. Rev.*, *57*(4), 193–216.

Solway, A., & Botvinick, M. M. (2012). Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates. *Psychol. Rev.*, *119*(1), 120–154.

Solway, A., & Botvinick, M. M. (2015). Evidence integration in model-based tree search. *Proc. Natl. Acad. Sci. U. S. A.*, *112*(37), 11708–11713.

Solway, A., Diuk, C., Córdova, N., Yee, D., Barto, A. G., Niv, Y., & Botvinick, M. M. (2014). Optimal behavioral hierarchy. *PLoS Comput. Biol.*, *10*(8), e1003779.

Sperber, D. (2006). Why a deep understanding of cultural evolution is incompatible with shallow psychology. In N. J. Enfield & Stephen C. Levinson (Ed.), *Roots of human sociality* (pp. 431–449). Routledge.

Strachan, J., Curioni, A., Constable, M., Knoblich, G., & Charbonneau, M. (2020). A methodology for distinguishing copying and reconstruction in cultural transmission episodes. In Denison, Stephanie and Mack, Michael and Xu, Yang and Armstrong, Blair C. (Ed.), *Proceedings of the 42nd annual conference of the cognitive science society.*

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

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, *364*(1528), 2405–2415.

Thorndike, E. L. (1932). *The fundamentals of learning.* Teachers College Bureau of Publications.

Tolman, E. C. (1948). *Cognitive maps in rats and men* (Vol. 55) (No. 4).

Tomasello, M. (1996). Do apes ape. *Social learning in animals: The roots of culture*, 319–346.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.*, *28*(5), 675–91; discussion 691–735.

Tomasello, M., Davis-Dasilva, M., Camak, L., & Bard, K. (1987). Observational learning of tool-use by young chimpanzees. *Hum. Evol.*, *2*(2), 175–183.

Vélez, N., & Gweon, H. (2021). Learning from other minds: an optimistic critique of reinforcement learning models of social learning. *Current Opinion in Behavioral Sciences*, *38*, 110–115.

Vikbladh, O. M., Meager, M. R., King, J., Blackmon, K., Devinsky, O., Shohamy, D., ... Daw, N. D. (2019). Hippocampal contributions to Model-Based planning and spatial memory. *Neuron*, *102*(3), 683–693.e4.

Whiten, A., & Ham, R. (1992). Kingdom: reappraisal of a century of research. *Adv. Stud. Behav.*, *21*, 239.

Wu, C. M., Schulz, E., Gerbaulet, K., Pleskac, T. J., & Speekenbrink, M. (2019). Under pressure: The influence of time limits on human exploration. In A. K. Goel, C. M. Seifert, & C. C. Freksa (Eds.), *Proceedings of the 41st annual conference of the cognitive science society* (pp. 1219–1225).

Zaki, J., Schirmer, J., & Mitchell, J. P. (2011). Social influence modulates the neural computation of value. *Psychol. Sci.*, *22*(7), 894–900.