nature human behaviour

PERSPECTIVE

https://doi.org/10.1038/s41562-019-0590-x

Resolving uncertainty in a social world

Oriel FeldmanHall* and Amitai Shenhav

Consider the range of social behaviours we engage in every day. In each case, there are a multitude of unknowns, reflecting the many sources of uncertainty inherent to social inference. We describe how uncertainty manifests in social environments (the thoughts and intentions of others are largely hidden, making it difficult to predict a person's behaviour) and why people are motivated to reduce the aversive feelings generated by uncertainty. We propose a three-part model whereby social uncertainty is initially reduced through automatic modes of inference (such as impression formation) before more control-demanding modes of inference (such as perspective-taking) are deployed to narrow one's predictions even more. Finally, social uncertainty is attenuated further through learning processes that update these predictions based on new information. Our framework integrates research across fields to offer an account of the mechanisms motivating social cognition and action, laying the groundwork for future experiments that can illuminate the impact of uncertainty on social cognition.

We are often faced with decisions that require us to consider our effect on others, as well as their effect on us. To date, research exploring social decision-making has focused on the vital role of reward and punishment in guiding choice. A large body of work now demonstrates that these decisions are driven by the same reward-related computations and neural circuitry as non-social decisions[1,2]. However, there is another critical—and equally potent—motivator of social behaviour: the desire to reduce uncertainty[3–6].

Interacting with others is one of the most inherently uncertain acts humans embark on. There are a multitude of unknowns, whether it is choosing how to express ourselves[7], who to confide in[8], how reliable an individual is[9,10] or whether to engage in risky behaviour with our peers[11,12]. The information we bring to bear on each of these decisions—from the expected trustworthiness or competence of others[8,13] to their anticipated reactions to an off-colour joke—relies on uncertain estimates[14]. It is therefore critical to our productivity, wellbeing and ultimately our survival as social beings to constantly estimate these uncertainties and find ways to reduce them[15].

A number of areas of social psychological research have illustrated that uncertainty in social environments is pervasive and aversive (Box 1). This research has further demonstrated that specific cognitive processes can reduce this uncertainty[3,16–18], including trying to identify which categories another person belongs to (for example, friend or foe). A largely separate research domain has further advanced our understanding of how people generate predictions regarding the likelihood of possible future events and has developed widely used quantitative approaches for estimating one's uncertainty[19]. We offer a framework that combines these two approaches to better explain the myriad influences of uncertainty on how we think and act in social environments. By addressing outstanding questions about how humans interact with such uncertainty—including (i) what factors give rise to social uncertainty, (ii) how people estimate and experience this uncertainty and (iii) what cognitive tools they use to resolve it—we hope to offer a path forward for theoretical and experimental work on the human relationship with social uncertainty.

What are the sources of social uncertainty?

It is now well-established that our brains function in part to generate predictions about potential future states, actions and outcomes[20–22]. Uncertainty describes the precision with which a prediction can be generated based on the available information[15,23]. We are therefore capable of being uncertain about everything our brain attempts to predict[15,23–25], be it features of stimuli (perceptual uncertainty), rewards or punishments that can be obtained (outcome uncertainty), actions to be selected (action uncertainty) or the manner in which those actions will be executed (motor uncertainty; see Box 2). Importantly, different sources of uncertainty build on one another, such that our uncertainty about a stimulus can increase our uncertainty about the potential outcomes it predicts, which in turn can increase our uncertainty about the best action to take[23,25,26].

From this perspective, social environments are particularly rife with uncertainty (Box 3). When interacting with others, each of our uncertainties about our own future states and actions is further compounded by the fact that we are often also uncertain about who these individuals are (their identities, characters and motives are largely hidden) and how they might choose to act in a given moment[27]. For instance, when we interview for a job, we may be uncertain about the interviewer's personality, the culture of the company and the skills most valued for the position, which will in turn make us uncertain about how to act to make the best impression. At any given moment, uncertainties about particular characteristics of the interviewer, such as his sense of humour, will increase our uncertainty about how to best answer a question and the kind of immediate feedback we will get as a result.

The degree of uncertainty for each of these predictions can be estimated in a variety of ways[23], including in terms of the variance of the predictions themselves (known as risk when predicting possible rewards[28,29]) and the variance in the probabilities assigned to those predictions (for example, the width[30] or the entropy of the probability distribution; see Box 3 [15,31]). While these measures place different weight on the expected outcome magnitudes versus the expected outcome likelihoods, they have a common feature: the more predictions a stimulus evokes—and thus, the more outcomes perceived to be likely—the greater the uncertainty, especially when many of those predictions are perceived to have similar likelihoods. Conversely, the higher the likelihood we place on a single prediction, the more certain we are.

Social uncertainty therefore increases with the number of plausible predictions we can generate about another person, including possible traits such as their warmth, trustworthiness and competence

Department of Cognitive, Linguistic, Psychological Sciences, Brown University, Providence, Rhode Island, USA. *e-mail: Oriel.feldmanhall@brown.edu

Box 1 | Foundations of research into social uncertainty

Our proposed framework builds on a rich psychological literature demonstrating that uncertainty can motivate certain social processes. For example, early work in the field of social psychology suggested that uncertainty motivates people to form quick, automatic initial categorizations, i.e., impressions or stereotypical expectations of other individuals. This idea was formalized in the Continuum of Impression-Formation model[3,56], which argued that an initial impression is integrated with subsequent information acquired about the person to either confirm the person's category or recategorize the individual into a new category. Following on this influential model, other researchers argued for the importance of documenting the epistemic motivations governing why an individual would make use of these processes[170]. For example, one theory posited that people need closure or desire structure to minimize ambiguity when a situation is sufficiently uncertain because they want to perceive their world in 'clear-cut' and unambiguous terms[171–173]. According to the Need for Closure model, people are highly motivated to 'seize' on information and then 'freeze' it into a lasting judgment, thereby providing cognitive closure and eliminating the need to collect more information about the world in the future[16,174]. Other theories have proposed that social uncertainty serves to motivate a range of behaviours, including those that maintain internal consistency[175], preserve a social identity[176,177], and enable affiliation or companionship[178]. Collectively, this body of research illustrates the specific ways in which social uncertainty helps shape behaviour and cognition, providing insight into the potential epistemic motives underlying this process.

More recently, some literatures have discussed uncertainty's role as a distinguishing feature between social and non-social environments[179], while others have highlighted the centrality of prediction to specific processes within social cognition[70,180,181]. This work has begun to describe how computational models can be used to simulate cognitive processes involved in social inference (such as theory of mind[70,182]). Such modelling is a critical step in the evolution of research on social cognition because it forces researchers to commit to a particular formalization of their theory within a common language, increasing the precision of those theories and the predictions they make. These models—and in particular the Bayesian and information theoretic frameworks discussed here—are therefore particularly beneficial when comparing theories of complex and unwieldly psychological processes (for example, inferring what another thinks), which can provide the quantitative groundwork for further research linking social uncertainty with social inference and learning[19,23,39].

Box 2 | Glossary

- **Social uncertainty**: The degree to which a person's uncertainty about (i.e., inability to precisely predict) their own future states and actions depends on their uncertainty about the states and actions of others.
- **Non-social uncertainty**: The degree to which a person's uncertainty is driven by predictions that do not primarily depend on how another person thinks, feels or acts. Non-social uncertainty encompasses the residual uncertainty one would experience in their daily life if the thoughts and behaviours of others were completely predictable.
- **Bayesian inference**: Generating predictions about potential states of the world by weighing the probabilities (likelihood) of new evidence against one's a priori beliefs (priors) to form an updated set of predictions (posterior).
- **Prior**: A distribution of probabilities that define one's belief state in the absence of (before obtaining) new evidence.
- **Entropy**: An information-theoretic measure of uncertainty based on a set of known event probabilities. Entropy is highest when these probabilities are all equal and decreases as a subset of events become more likely than others.
- **Perceptual uncertainty**: Uncertainty about features of stimuli in one's environment (such as shapes or colours). For social stimuli, these may include labels (for example, names or group affiliations), facial expressions (for example, smiling or frowning), and social norms associated with environmental settings (for example, church or bar).
- **Action uncertainty**: Uncertainty about which action to take in one's current state (for example, turn left or right). For social stimuli, these actions include possible verbal or nonverbal communications, aggressive or affiliative actions and transactions (for example, lending money).
- **Outcome uncertainty**: Uncertainty about what kinds of rewards or punishments one could receive (for example, amounts of monetary gain or loss). For social stimuli, these outcomes can be concrete (such as money gifted or stolen) or abstract (for example, appreciation or approbation), and they can affect future rather than immediate states (for example, obtaining someone's trust or having your reputation impugned).
- **Impression formation**: A rapid evaluation of a person's physical features to help determine group membership.
- **Theory of mind**: The ability to infer another person's mental states (thoughts, perceptions, motivations).
- **Affect-sharing**: Attempting to experience another person's current emotional state.
- **Explore–exploit dilemma**: The tension between choosing an option that has a known (certain) reward distribution (for example, one's default menu item; exploiting) or choosing an option with an uncertain reward distribution (for example, new item on the menu; exploring) to collect new information and reduce one's uncertainty. In social environments, these can manifest in terms of choices to interact with close friends or unknown acquaintances.

(Box 3). Even with a reasonably good estimate of another person's more stable features (as with for a close friend or long-term acquaintance), any given interaction still carries uncertainties regarding that person's intentions and motivational state in that moment, which may diverge from what they are expressing[32,33]. All of these factors bear on the particular words or actions you select when deciding how to interact with that person: will your joke produce a laugh, an eye-roll or a look of horror and disbelief? These forms of uncertainty manifest to varying degrees for every person we encounter, making uncertainty endemic and unavoidable in the social world.

How do people react to social uncertainty?

Substantial research on non-social decision-making has documented the ways in which individuals, groups and organizations respond to the forms of uncertainty outlined above[15,23,34–37]. This work has enumerated how uncertainty is estimated and how it is used when integrating over different sources of information[25,38,39], evaluating potential actions[40] and updating expectations based on feedback[41,42]. In addition to demonstrating the many ways in which people navigate the uncertainty in their environment, this research reveals that uncertainty also tends to trigger negative affective reactions (such as anxiety)[15,43–45]. People typically experience uncertainty as aversive, and this provides them with an additional motivation to reduce it, independent of whether it is adaptive to do so[46].

Box 3 | Social uncertainty: a formal description

We have described a quantitative framework by which social information is incorporated into our predictions about our environment (Figs. 1 and 2), but a critical question is how these predictions form the basis of one's experience of uncertainty. How humans react to uncertainty has been documented through various integrative accounts[3,15,23], including a recent framework operationalizing non-social uncertainty in terms of entropy[15,31,166,183], an information-theoretic construct that measures uncertainty in terms of the probabilities of discrete outcomes occurring, $\Pr(x)$:

$$\text{Shannon's entropy} = \mathrm{H}(\mathrm{x}) = -\sum_j \Pr(x_j) \cdot \log_2(\Pr(x_j))$$

Entropy is lowest when a single value of x (for example, a particular outcome) is nearly certain, and it increases as there are more values of x (for example, many possible outcomes) that each have increasingly similar likelihoods. Changes in the width of a probability distribution as one updates their beliefs about their environment (Fig. 2) will therefore tend to generate correlated changes in the entropy of that distribution. We leverage this framework to offer a more formal account of social uncertainty and its sources and to capture the stark differences in uncertainty presented by social versus non-social settings outlined in the main text.

We begin by assuming that at any given time, there is a distribution of actions (a) an individual may take given their current state (s_y). We can describe their uncertainty over those possible actions in terms of the conditional entropy of this distribution (the entropy over a set of conditional probabilities):

$$\text{Total uncertainty (nonsocial + social)}$$
$$= \mathrm{H}(a|s_y) = -\sum_j \Pr(a_j|s_y) \cdot \log_2(\Pr(a_j|s_y)) \cdot \Pr(s_y)$$

When a person is preparing to brush their teeth, for instance, certain actions have a much greater likelihood than others (low entropy). When they are instead preparing to write a manuscript, many more actions are likely (high entropy). Given that our actions can determine our future states, uncertainty regarding potential actions propagates to uncertainty over those future states, and this overall uncertainty results in an increasingly negative affective state.

When another individual (i_z) is present (for example, an interviewer or a colleague), these uncertainties typically multiply. Given that one's own actions depend on how the other person might act (or react) in a given situation, uncertainty regarding one's own actions (for example, what to say or how to position one's face and body) is tied to the ability to predict another's actions in the current state. Predicting the other person's actions requires further deconstructing that person into uncertain attributes, for instance, what kind of person i_z is, Pr(trait t), where possible traits $t \in$ {trustworthy, kind, competent, …}; what are their current intentions, Pr(goal g), where $g \in$ {asserting dominance, networking, making new friends, …}; and how they are feeling, Pr(emotion e), where $e \in$ {happy, angry, disappointed, …}. Thus, our own uncertainty about how to act in a social setting is influenced by how uncertain we are about each of these social attributes:

$$\text{Social uncertainty} = \mathrm{H}(a \mid s_y, i_z)$$
$$= \mathrm{H}(a \mid s_y, \textit{trait}, \textit{goal}, \textit{emotion}, \ldots)$$
$$= -\sum_j \sum_t \sum_g \sum_e \cdots$$
$$\Pr(a|s_y, \textit{trait}_t, \textit{goal}_g, \textit{emotion}_e, \ldots)$$
$$\cdot \log_2(\Pr(a_j|s_y, \textit{trait}_t, \textit{goal}_g, \textit{emotion}_e, \ldots))$$
$$\cdot \Pr(s_y, \textit{trait}_t, \textit{goal}_g, \textit{emotion}_e, \ldots)$$

As the possible values of each of these variables increases, so does our uncertainty about how that person might act (for example, are they preparing to compliment or reprimand you), what our best course of action is in that moment and how another might react to our action. Additionally, the potential outcomes that could occur through another person's involvement (for example, what the most positive or negative outcome could be: $\max(|\mathrm{Value}(\textit{outcome}|s_y,i)|)$ serve as a marker of motivational relevance that can enhance our sense of uncertainty. Critically, these uncertainties exist over and above the non-social sources of uncertainty described above; uncertainties that arise over manuscript writing in isolation will be modulated by uncertainties over potential readers ($a|s_y,i$).

However, as illustrated in Fig. 2, these uncertainties decrease with automatic inference (for example, narrowing a person's likely traits based on impression formation), controlled inference (for example, narrowing a person's likely intentions based on perspective-taking) and learning. As these predictions are updated through Bayesian inference (for example, as the distribution of possible traits and intentions narrows), the relative entropy between the posterior and the prior distribution provides an estimate of the information gained by a given update.

There is substantial evidence that uncertainty generates aversive reactions in both the non-social domain[15,47,48] and the social domain[49,50]. However, a key prediction of our model is that the main difference between these two is the amount of uncertainty generated. Social stimuli are inherently more unpredictable (and thus more difficult to predict) than their non-social analogs (Box 3): not only are the causes of other peoples' behaviours uncertain (for example, their motives are not directly observable), they are also dynamic and constantly evolving[51]. As a result, uncertainty about everything from choice of attire to choice of emotional expression will be magnified when considering those you might interact with, which increases the potential for negative affect. Moreover, the worst consequences of a misstep in a social exchange (for example, making an inappropriate comment) can be devastating and long-lasting, including the potential for damaged interpersonal relationships, social isolation, and persistent loneliness[52,53].

How do people resolve social uncertainty? A model of social uncertainty

Given the manifold sources of aversive uncertainty provided by social environments, it is no surprise that social stimuli motivate human social cognition and behaviour to the extent they do. This motivation to reduce uncertainty offers a unique window into the kinds of cognitive processes that unfold in these social settings. Our framework predicts that people are motivated to think about and act towards others in ways that reduce their own uncertainty and the attendant negative affect (Figs. 1 and 2). The objective is therefore to narrow the range of possible outcomes another person

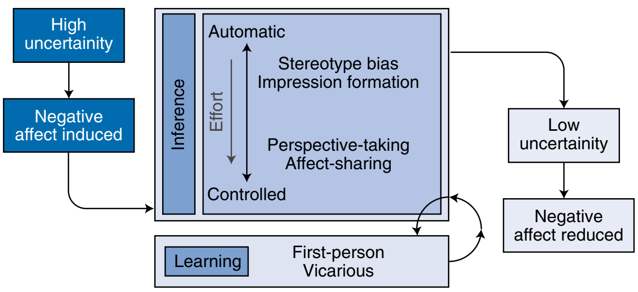

Fig. 1 | Model for how humans resolve social uncertainty. Social environments evoke high degrees of uncertainty. Because uncertainty is aversive (particularly when it involves a salient and motivationally relevant outcome), most people have an intrinsic desire to reduce this uncertainty. This motivates three interrelated mechanisms. More-automatic inferential processes quickly narrow one's predictions using past knowledge and contextual cues, while more controlled inferential processes further hone these predictions through a more effortful search over an internal model of the other person's thoughts and feelings. These and other forms of social inference fall along a continuum of automaticity, with processes like perspective-taking and affect-sharing varying in their demands on cognitive control (and associated effort costs) depending on the individual and context. Finally, learning processes update predictions based on feedback. Together, these processes serve to lower uncertainty and the associated feelings of negative affect.

portends, in turn making one's own future states more predictable (Fig. 2 and Box 3).

Our account also delineates the processes by which people try to reduce their uncertainty and the far-reaching consequences this has for interactions at the interpersonal and societal level. In particular, we propose that social uncertainty motivates three forms of interrelated mechanisms that can help reduce it: relatively automatic inferential processes that quickly narrow one's predictions using past knowledge and contextual cues, more control-demanding processes that further hone these predictions through an effortful search over an internal model of the other person's thoughts and feelings, and learning processes that update one's predictions based on feedback (Fig. 1). We further describe how these processes can be readily aligned with and operationalized by a Bayesian framework[25,26], according to which uncertainty is determined by a combination of the available evidence and learned priors.

Automatic inference. For every individual that we encounter, several cognitive mechanisms—which are unique to the social domain—immediately come online to shape social perceptions and form impressions[54,55]. This initial process of impression formation simultaneously evaluates the person's physical features (such as their clothes, skin colour, hair style[56]) and the social norms and moral rules inherent to the environment (for example, whether the situation occurs in a church or at a bar), determining, for instance, likely categories of occupation and group membership[3,57–59]. The resulting predictions (for example, regarding group membership) serve as split-second judgments we make about what this person might be like, including whether they are trustworthy, competent, kind or threatening[60–63]. These automatic early impressions are largely unaffected by other ongoing processes[64,65], and they form strong priors based on the features of the person and their environment, collectively constraining our predictions about that person[66] (Fig. 2). Consistent with instances of Bayesian inference in other domains (such as perceptual decision-making[25]), these priors can play a particularly outsize role in initial evaluations of others when we have little information about them. For example, when uncertainty in the environment is particularly high, people appeal to established social norms, such as fairness and equity considerations, as a reference for how to engage with others[67,68].

These initial impressions thus help to constrain our predictions about who a person is, what drives them, what they are capable of and how they might act (Fig. 2). In so doing, the process of impression formation necessarily prunes our prediction space, making others more predictable and making us less uncertain about how to act in their presence. Returning to our interview example,

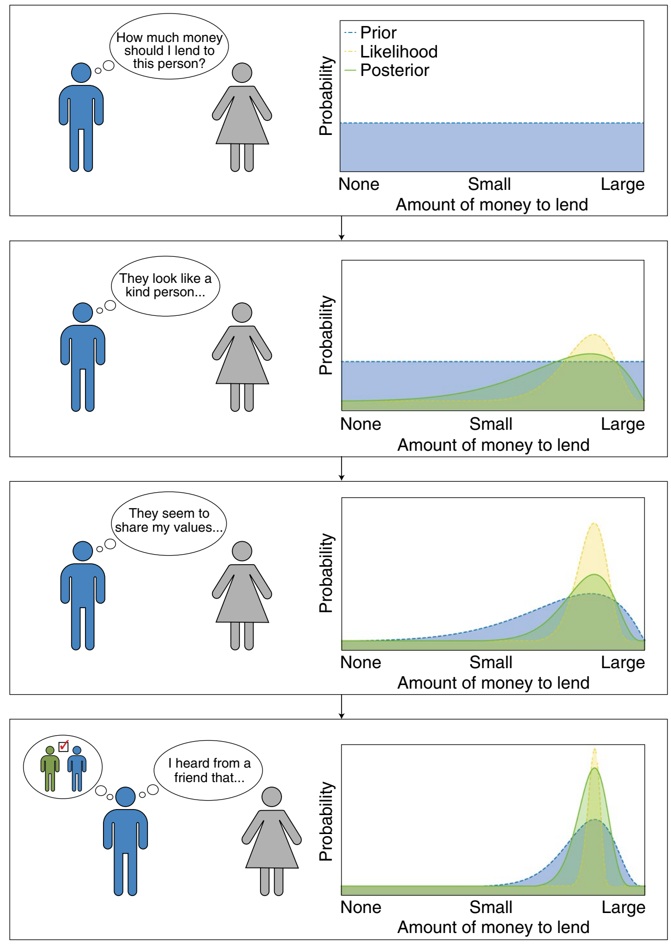

Fig. 2 | Iterative reduction of social uncertainty through inference and learning. Here we imagine an agent who is confronted by an unfamiliar person asking for a loan. When deciding how much money to lend, the agent initially assigns each possible action (for example, amount of money lent) equal likelihoods. The corresponding flat distribution of predicted actions (i.e., priors, denoted in blue) represents maximal uncertainty (Box 3). (Note, however, that there are many cases in which this initial distribution may not be flat; for example, when there are priors associated with the person or their social category.) As the agent engages in automatic inferential processes (for example, impression formation), the associations generated may support a narrower set of actions (such as lending a relatively large sum, the likelihood function; yellow); this information will be weighed against the agent's priors to determine how likely the agent would be to offer different sums (the posterior; green). When controlled inferential processes come online (sharing in the emotional experience of the other), the distribution of possible actions continues to narrow in on a more specific amount to lend. Finally, the agent can acquire new information about the prospective borrower (in this case, hearing from a friend that that this individual is highly trustworthy; i.e., vicarious learning), narrowing the prediction space further and increasing the agent's certainty about which amount to lend.

Box 4 | Is reducing uncertainty always beneficial?

Patient populations such as those with anxiety or autism spectrum disorders provide clear examples of the benefits humans accrue from resolving social uncertainty and the distress it may cause in their lives. However, the approaches we take to resolving uncertainty can also come at a cost to ourselves and to society at large. For instance, since we are most certain when we have the narrowest set of predictions about what will happen, one clear way of decreasing our uncertainty is to develop a habit[30,184,185]. However, these habits can become maladaptive when they manifest as compulsions in obsessive–compulsive disorder[186,187] or other forms of stereotyped behaviours in autism spectrum disorders[153,188] (see also ref. [189]). Though social rituals—which are believed to draw on similar mechanisms as habits[186]—serve a number of adaptive purposes (see main text), they also have the potential to be highly maladaptive (for example, hazing practices[190]). In each of these cases, greater certainty trades off against greater flexibility (i.e., goal-directedness)[30,184,191].

A similar trade-off occurs in the case of stereotypes, which are efficient at constraining our predictions of another's behaviour but often lead to inaccurate inferences that fail to be disconfirmed. As has been observed with other forms of strong cognitive priors[192–194], these social priors are likely to generate a relatively narrow distribution of predictions and may serve an adaptive purpose in guiding snap judgments about who to approach or avoid, but are on their own insufficient and potentially misleading when one's goal is to form accurate predictions of another's behaviour. More generally, when we anchor our inferences about another person on ourselves (or a group prototype), research shows that we do not always invest the time and effort to appropriately adjust from those initial starting points[126,195,196]. As a result, our predictions about that person become systematically biased towards our own egocentric anchors. Such failures to sufficiently adjust from our anchors can lead to negative behavioural patterns, such as favouring in-group members despite information that one should not, or even depersonalizing and stigmatizing others[197]. When under stress or cognitive load—a frequent feature of daily life—these predictions can become even more biased toward the initial anchor, since the adjustments are cut off early in the tuning process[112,198].

a quick glimpse of the interviewer will allow us to infer how they carry themselves, which in turn will constrain what we choose to say. Unlike other forms of social inference (discussed below), this process of updating our predictions about another person comes with few cognitive costs; it is triggered automatically when viewing others, in the sense that it is rapid and largely unaffected by other ongoing cognitive processes[64,65], and therefore requires little effort[69]. A cognitive advantage of rapidly classifying people into relevant categories is that we do not spend unnecessary resources trying to predict the behaviours of those we are somewhat certain about. This provides a relatively effortless method of reducing the negative affect associated with uncertainty—a method, however, that can also carry attendant costs for society at large (for example, stereotyping; Box 4).

Controlled inference. Automatic forms of inference constrain our priors regarding what to expect from another person, but they do so in a relatively coarse manner, reflecting a rapid integration of our prior knowledge with the environmental information most readily available. They identify the categories to which a person likely belongs but leave many possibilities for how this particular individual—as an exemplar of those categories—will actually behave.

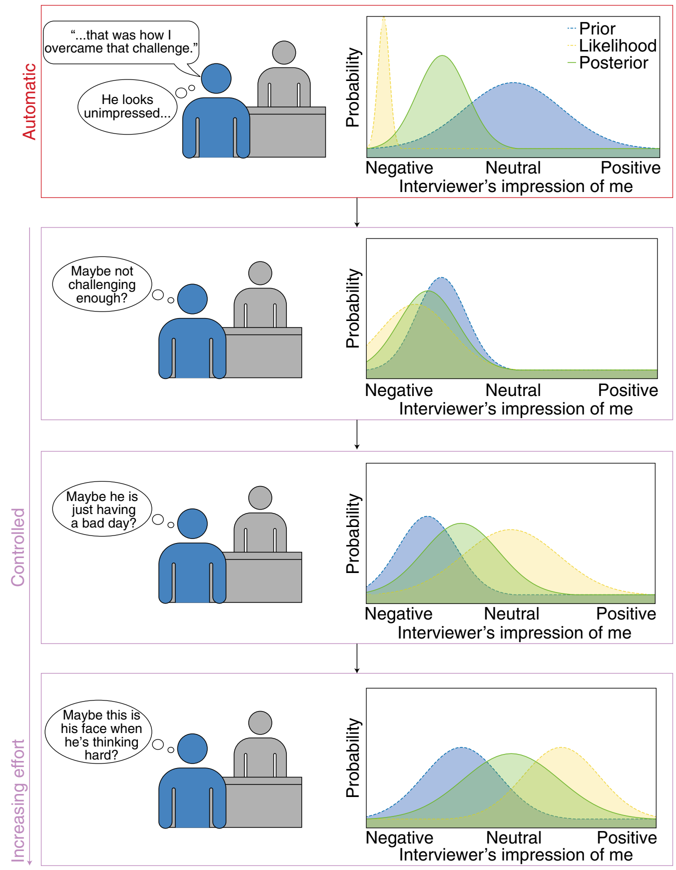

Fig. 3 | The unfolding of automatic and controlled components of social inference. When interacting with another person (for example, an interviewer), we dynamically adjust our estimates of what that person is thinking through processes that are more automatic or more controlled. For instance, if the interviewer grimaces after we give an answer, we may automatically encode this as strong evidence of a negative impression, weighing this against our prior (for example, of a neutral impression). However, the resulting estimate can be adjusted further by engaging in more controlled forms of inference (such as perspective-taking). This might lead us to consider other factors that would help contextualize the interviewer's facial expression (such as how they might have perceived our answer, their overall mood, and how they express their emotions). Such inferences allow us to iteratively update our predictions about another person based primarily on endogenous information (versus explicit feedback). However, each of these sequential inferences incurs a cognitive cost, resulting in a bias towards relying more on our automatic estimates (cf. anchors).

Fortunately, we have several additional tools at our disposal, including the ability to generate an internal model of that person's thoughts and feelings, which can be subsequently used to further reduce social uncertainty (Fig. 3).

From a Bayesian perspective[37], controlled inference can be viewed as the process by which our early impressions are updated in light of incoming information[42]. For example, by engaging in perspective-taking, we try to consider the current environment from another's point of view to better predict their future actions[70,71] and assume that they have different priors or access to different types of evidence[72]. Recent work has applied such an approach to modelling behaviour in a competitive social game, revealing that variations of Bayesian inference can capture the sophistication with which humans and non-human primates use theory of mind to infer the intentions of others[73,74]. We can also engage in affect-sharing to try and place ourselves in that person's current emotional state[75], which would further narrow our predictions of their possible actions or reactions[76,77]. These inference processes—which are honed over development[78,79]—often require greater cognitive control and are therefore more computationally expensive and effortful than automatic inferences[80,81], so much so that people will at times forego potential rewards to avoid the effort of sharing in another's affective state[82].

While automaticity and control provide useful reference points in considering the taxonomy of inferential processes, previous research has established that these are best understood as forming a continuum rather than a dichotomy[83,84]. Processes can become more or less automatic depending on their context and how well-practiced they are[65,85]. We likewise expect these social inferences to become more automatic (i.e., less control-demanding) as an individual further develops these skills, for instance with age[86–88]. Along the same lines, not all methods of inference are monolithic in their control demands: there are forms of perspective-taking and affect-sharing that are believed to be more automatic and may transpire in parallel with the more effortful processes described above[89–92]. For instance, processes like emotional mimicry and contagion[93,94] may come online relatively automatically, especially in conditions when the target is physically or socially close.

Social learning. Together, both automatic-controlled inferences constrain our predictions about another's personality and intentions, thereby reducing our uncertainties regarding how that person might think and behave (Fig. 2). These mechanisms rely on learned predictions regarding how a person might act given, for instance, the social group they belong to. Those predictions can in turn be updated based on at least two forms of feedback[95]. First, we can directly observe how a person or group behaves in different situations or in response to actions that we take (for example, indications that the interviewer approves of our answers to his question; Fig. 3)[13,96]. Second, we can obtain second-hand information about an individual based on the experiences of others (vicarious learning: for example, you heard that the interviewer is sexist)[97–100]. In either case, we update our predictions about that individual or group based on a weighted combination of the new evidence and our prior predictions (Figs. 2 and 3). Depending on whether the new information is consistent or inconsistent with those priors—and whether it accords with or deviates from other strong learned social associations (cf. blocking[101,102])—it can serve to narrow or broaden our distribution of subsequent expectations (i.e., posteriors) about an individual's motives and likely actions, thereby decreasing or increasing our uncertainty about that individual.

In addition to helping reduce uncertainty, it is likely that social learning is also modulated by uncertainty, with the rate of learning increasing with greater uncertainty[41,103,104]. For example, one strength of repeated economic games is that they allow individuals to learn over time about the social value of another person based on their behavioural patterns[105,106] or to glean information about the value of environmental stimuli from others (i.e., advice-giving[98,107]). In either case, learning rates are likely highest in the beginning of the experiment when there is maximal uncertainty about how others will act[98,108,109], suggesting that we learn about our social worlds more quickly when uncertainty is great.

Our model therefore suggests that multiple forms of Bayesian updating can occur during inferential processing (automatic and controlled) and when learning about others using feedback. In both cases, we start with weaker predictions (flatter probability distributions), and these predictions are sharpened with additional evidence (Figs. 2 and 3). In the case of inference, we use a previously learned model to sharpen our predictions based on initial impressions (automatic inference), as well as additional information we generate about the person and their context (more controlled inference). In the case of learning, we update those predictive models based on feedback that we, or others, observe. While some of the same learning processes can also be captured by traditional reinforcement learning (for example, *Q*-learning) algorithms[110], a strength of a Bayesian approach is that predictions are updated based on internal and external sources of information about the environment in proportion to one's uncertainty about those predictions. Importantly, though, while our account assumes that social inference follows the general principles of Bayesian updating (i.e., uncertainty-weighted updating of priors given observed distributions of possible outcomes), it does not require that people behave in a Bayes-optimal fashion. Rather, it is likely that people also use strategies that approximate Bayesian inference, for instance by basing a given judgment only on their priors (interviewers are notoriously critical), or only on the current evidence (the likelihood; for example, the interviewer is frowning)[111] or by engaging in a process of sequential hypothesis testing that initiates with the prior[112–114].

When do we reduce social uncertainty?

These interrelated mechanisms demonstrate the array of tools people deploy to reduce social uncertainty and the accompanying aversive feelings. However, the extent to which one engages these mechanisms depends on several factors, including the importance of the other person and the availability of our cognitive resources. Humans are only motivated to reduce social uncertainty if it is something they care about reducing (for example, if the situation or person can directly influence our ability to achieve our goals)[15,115–117]. For example, we may be most likely to engage in empathic perspective taking for people with whom we want to continue to engage or who we think will be important to our future interests. Those whose actions have the ability to affect us—whether in the immediate future or over the long-term (for example, a boss, family member, close friend)—typically hold greater motivational salience than those who do not have the power to affect our future states. Uncertainty about the actions or intentions of such people has the greatest bearing on our own future states and actions, which in turn increases our desire to engage in processes that minimize the attendant social uncertainty.

What shortcuts do we use for reducing social uncertainty?

We are economical in how we allocate our cognitive resources when resolving uncertainty[80], relying on mental shortcuts to assign probabilities to particular events[4,80,118–120]. One of the most common examples of these shortcuts is the use of informative anchors as a starting point when making an inference[121]. For example, to predict another's traits we often anchor on the social categories a person most prototypically encapsulates[80]. Such anchors may reflect priors that are relevant to the specific person and situation (such as stereotypes; Box 4), facilitating our ability to generate accurate predictions about their behaviour, or they may reflect a convenient and readily accessible heuristic that can serve the same purpose[121,122]. Adjusting away from these anchors is computationally costly because it requires an iterative process of hypothesis testing[112]. Those adjustments take time and mental effort[112,123–125], and people therefore adjust less and are more biased towards the initial anchor when they are cognitively taxed[126]. As a result, anchors related to general properties of an individual form a path of least resistance that must be overcome in order to generate more precise inferences about them.

We also use another efficient and salient anchor when predicting how someone else will behave: ourselves[123]. Research on social comparisons illustrates that, to generate quick and economical estimates of how likely another's preferences align with our own, we routinely assess whether another person is similar to us[115]. The greater the similarity, the more we can use our own thoughts and behaviours as a proxy for theirs, which aids in resolving uncertainty about their future behaviour. Conversely, dissimilarity between ourselves and another increases uncertainty[127], since the number of explanations for another's actions is much larger[14,128]. Ultimately, the nature of these comparisons necessarily biases our subsequent choices about how and when to engage with specific individuals.

What are the consequences of reducing social uncertainty?

According to our model, one of the ultimate goals of reducing uncertainty about others' intentions and behaviour is to minimize the aversive affect associated with being uncertain about our own future states. This goal motivates people to use methods of inference and learning described above to reduce their uncertainty about others in their social environments. But it can also motivate people to avoid uncertain social situations, leading them to prefer spending time interacting with individuals who are more, rather than less, predictable (for example, members of similar social categories[129,130]). By avoiding these situations altogether, a person can prevent both the negative experiences of uncertainty and the cognitively taxing experience of reducing it. Indeed, those who are less tolerant of uncertainty avoid engaging with other people in social situations that are highly uncertain, for instance in economic games where they don't know whether the other players will reciprocate their trust[131].

At the same time, identifying with social groups helps to satisfy the reduction of uncertainty, since belonging to certain groups (ranging from clubs and teams to cultures and religions) offers moral codes and social mores about how we should think, behave and perceive others[132]. Indeed, our account would predict that people should prefer to engage in specific activities that increase their certainty about how to act in a given social environment. One of the best examples of such an activity is a ritual, a set of behaviours that a group performs in a stereotyped fashion[133,134]. Whether in a church, stadium, or at home, rituals offer an individual greater certainty about the kinds of behaviours group members will be expected to perform within a given setting. By helping us manage uncertainty, even the most arbitrary rituals can help relieve distress and improve performance on ongoing tasks[135,136]. Indeed, the sheer array of rituals people engage in over their lifetime and across generations provides some insight into just how aversive people find social uncertainty and the lengths they will go to reduce it.

While people are generally averse to uncertainty—and social settings magnify uncertainty by introducing dynamic elements that influence our future states in ways that can be difficult to predict—these reactions vary considerably across individuals[131,137] and contexts[138–140]. Previous research has demonstrated that trait variability in people's negative reactions to uncertainty (for example, individual differences in intolerance of uncertainty) correlates with anxiety-related impairments across several clinical disorders, including social anxiety[45,141–146] and major depression[147,148] (see also ref. [149]). For people with high intolerance of uncertainty, the increased uncertainty inherent to social settings—and subsequent ruminations over those uncertainties—may trigger particularly heightened levels of anxiety. In such cases, the enhanced aversion to uncertainty can inhibit action, which in turn may limit an individual's ability to gain information about the environment that could ultimately reduce the uncertainty.

Nevertheless, these anxieties can be attenuated by employing cognitive tools that reduce uncertainty[150]. There are, however, other populations that do not have access to the tools needed to reduce uncertainty. For instance, individuals who have difficulty inferring the intent of others or communicating their own intentions, such as those with autism spectrum disorders, are impaired at resolving social uncertainty[151,152]. More than their neurotypical peers, such populations experience the world as though a great many more outcomes are possible whenever another individual is present, resulting in a high degree of unresolved social uncertainty and associated negative affect. This unresolved uncertainty may in turn serve to facilitate forms of behavioural inflexibility that are also diagnostic of autism spectrum disorders[153] (Box 4).

Box 5 | Outstanding questions

- Social settings tend to evoke greater uncertainty than non-social settings. But do equivalent levels of uncertainty evoke greater negative affect in social settings than in non-social settings? More generally, what is the form of the relationship between uncertainty and negative affect? Does it increase linearly or nonlinearly, for instance, as an exponential, logarithmic or step function? Does intervention on these affective reactions (administering anxiolytics) diminish the motivation to reduce uncertainty?
- Which source(s) of social uncertainty produce the greatest negative affect? Are these aversive experiences best accounted for by uncertainty about actions, outcomes or the characteristics of others? For instance, is it the case that we are only averse to not knowing information about a stranger if we think that information will impact our future actions, as our current operationalization predicts?
- Similarly, what formalization of motivational relevance best explains when uncertainty is particularly potent? For instance, is motivational relevance determined by the value of the most extreme or the most likely outcome that could result from another person's involvement (for example, being offered a new job or getting fired from your current one)?
- Since controlled inferences are effortful, they require the motivation of uncertainty. An open question, however, is whether there are cases in which automatic inferences are motivated to the same extent. Or, in contrast, do automatic processes proceed independently of one's uncertainty? In a similar vein, can parameters of the learning process (such as learning rate) be modulated by uncertainty?
- How do people trade off the costs associated with being in a socially uncertain state (i.e., negative affect) against the costs associated with engaging in controlled inference to reduce their uncertainty (i.e., mental effort)? Given that certainty does not guarantee accuracy[199,200], do preferences for certainty also trade off against preferences for accuracy? Across individuals, does uncertainty avoidance positively correlate with effort avoidance and/or negatively correlate with accuracy bias?
- Is social uncertainty encoded by a system dedicated to processing social information, or by a domain-general system? Similarly, are signals of social uncertainty and social rewards generated by distinct or overlapping circuits?
- Do individuals with social anxiety only experience enhanced negative affect under uncertainty or are they also impaired at reducing uncertainty?
- Under what conditions do people seek out social environments with more rather than less uncertainty? Do individual differences in uncertainty seeking reflect positive feelings towards uncertainty itself or to the desire for information and/or for the resolution of aversive uncertainty?

Even within a given individual, there are contexts in which uncertainty is treated as appetitive rather than aversive (Box 5). For instance, the goal of reducing uncertainty can come into tension with the goal of avoiding aversive states that result from the resolution of uncertainty, as in the case of deciding whether to learn the outcome of a medical test or a teaching evaluation[154–156]. In such instances, the individual may prefer greater uncertainty to the certain knowledge of a potentially devastating outcome[157]. Conversely, when faced with the choice of exploiting an option with a known reward distribution (for example, having lunch with an old friend) versus exploring an option with an unknown distribution (for example, having lunch with a new colleague), people often choose to explore[158–160]. People also demonstrate a similar preference for novel over familiar options[161], again seeming to favour greater uncertainty when all else is equal. While such preferences might at first appear to reflect a positive association with uncertainty, it is also likely that these preferences are often motivated by a desire to collect additional information about the environment[162,163] or to simply reduce feelings of boredom or monotony[164,165]. Moreover, by engaging with an uncertain option, the individual's actions necessarily serve to reduce that uncertainty[20,166] and any aversive experiences associated with it—something that has been previously noted in research on curiosity (another phenomenon that appears to suggest a preference for uncertainty)[167–169]. Curiosity therefore serves to reduce uncertainty in the long term, at the cost of short-term engagement with uncertain options.

Conclusions

Despite substantial research into the sources, mechanisms and consequences of uncertainty within the non-social domain, less empirical work has examined the role uncertainty plays in motivating social behaviour and cognition. Our framework integrates research across several fields in order to offer an account of the mechanisms motivating social cognition and action. We highlight how pervasive uncertainty is in our social world: we are uncertain about the intentions and actions of each person we encounter, which leads us to be uncertain about our own future states and potential actions. This uncertainty generates aversive feelings, driving us to attempt to narrow the range of predicted actions that person might take, thus making our own future more predictable. We posit that three forms of interrelated mechanisms govern this process of resolving uncertainty: automatic inference, controlled inference and learning.

Our account also describes how uncertainty can systematically bias individuals, groups and societies towards heuristics and other forms of stereotyped behaviours that artificially resolve these uncertainties. While these can be efficient methods of reducing negative affect, such heuristics can also lead to inaccurate and systematically biased inferences about others. Finally, we account for the particular potency of social environments to act as triggers for aversive reactions, especially in clinical populations who are marked by an enhanced reactivity to uncertainty and/or an impaired ability to engage in inferential processes required to reduce this uncertainty. By highlighting uncertainty's role in motivating social behaviour, we hope our model helps provide a better understanding of when the human drive to reduce uncertainty leads us to behave in ways that are more—or less—beneficial to ourselves and those around us.

Received: 13 August 2018; Accepted: 15 March 2019;
Published online: 22 April 2019

References

1. Fehr, E. & Camerer, C. F. *Trends Cogn. Sci.* **11**, 419–427 (2007).
2. Ruff, C. C. & Fehr, E. *Nat. Rev. Neurosci.* **15**, 549–562 (2014).
3. Fiske, S. T. & Neuberg, S. L. *Adv. Exp. Soc. Psychol.* **23**, 1–74 (1990).
4. Festinger, L. *Hum. Relat.* **7**, 117–140 (1954).
5. Bernoulli, D. *Econometrica* **22**, 23–36 (1954).
6. Alchian, A. A. *J. Polit. Econ.* **58**, 211–221 (1950).
7. Camerer, C. *Behavioral Game Theory: Experiments in Strategic Interaction*. (Russell Sage Foundation; Princeton University Press, 2003).
8. Fehr, E., Fischbacher, U. & Kosfeld, M. *Am. Econ. Rev.* **95**, 346–351 (2005).
9. Berg, J., Dickhaut, J. & McCabe, K. *Games Econ. Behav.* **10**, 122–142 (1995).
10. Krueger, F. et al. *Proc. Natl Acad. Sci. USA* **104**, 20084–20089 (2007).
11. Bechara, A. *J. Gambl. Stud.* **19**, 23–51 (2003).
12. Albert, D., Chein, J. & Steinberg, L. *Curr. Dir. Psychol. Sci.* **22**, 114–120 (2013).
13. King-Casas, B. et al. *Science* **308**, 78–83 (2005).
14. Tarantola, T., Kumaran, D., Dayan, P. & De Martino, B. *Nat. Commun.* **8**, 817 (2017).
15. Hirsh, J. B., Mar, R. A. & Peterson, J. B. *Psychol. Rev.* **119**, 304–320 (2012).
16. Kruglanski, A. W. & Webster, D. M. *Psychol. Rev.* **103**, 263–283 (1996).
17. Neuberg, S. L. & Newsom, J. T. *J. Pers. Soc. Psychol.* **65**, 113–131 (1993).
18. Cacioppo, J. T. & Petty, R. E. *J. Pers. Soc. Psychol.* **42**, 116–131 (1982).
19. Griffiths, T. L., Kemp, C. & Tenenbaum, J. B. Bayesian models of cognition. in *The Cambridge Handbook of Computational Psychology* (ed. Sun, R.) 59–100 (2008).
20. Friston, K. *Nat. Rev. Neurosci.* **11**, 127–138 (2010).
21. Bar, M. *Trends Cogn. Sci.* **11**, 280–289 (2007).
22. Barrett, L. F. & Simmons, W. K. *Nat. Rev. Neurosci.* **16**, 419–429 (2015).
23. Bach, D. R. & Dolan, R. J. *Nat. Rev. Neurosci.* **13**, 572–586 (2012).
24. Kagan, J. *Rev. Gen. Psychol.* **13**, 290–301 (2009).
25. Körding, K. P. & Wolpert, D. M. *Trends Cogn. Sci.* **10**, 319–326 (2006).
26. Griffiths, T. L. & Tenenbaum, J. B. *J. Exp. Psychol. Gen.* **140**, 725–743 (2011).
27. Calabrese, R. & Berger, C. R. *Hum. Commun. Res.* **1**, 99–112 (1975).
28. Glimcher, P. W. *Cogn. Affect. Behav. Neurosci.* **8**, 348–354 (2008).
29. Knight, F.H. *Risk, Uncertainty and Profit*. (Courier Corporation, 2012).
30. Daw, N. D., Niv, Y. & Dayan, P. *Nat. Neurosci.* **8**, 1704–1711 (2005).
31. Shannon, C. E. *Bell System Tech. J.* **27**, 379–423 (1948).
32. Apperly, I. A., Back, E., Samson, D. & France, L. *Cognition* **106**, 1093–1108 (2008).
33. Decety, J., Smith, K. E., Norman, G. J. & Halpern, J. *World Psychiatry* **13**, 233–237 (2014).
34. Milliken, F. P. *Acad. Manage. Rev.* **12**, 133–143 (1987).
35. Camerer, C. & Weber, M. *J. Risk Uncertain.* **5**, 325–370 (1992).
36. Downey, H. K. & Slocum, J. W. *Acad. Manage. J.* **18**, 562–578 (1975).
37. Gilboa, I. & Schmeidler, D. *J. Math. Econ.* **18**, 141–153 (1989).
38. Wolpert, D. M. & Landy, M. S. *Curr. Opin. Neurobiol.* **22**, 996–1003 (2012).
39. Summerfield, C. & Tsetsos, K. *Trends Cogn. Sci.* **19**, 27–34 (2015).
40. Frank, M. J. et al. *J. Neurosci.* **35**, 485–494 (2015).
41. Yu, A. J. & Dayan, P. *Neuron* **46**, 681–692 (2005).
42. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. *Nat. Neurosci.* **10**, 1214–1221 (2007).
43. Gray, J.A. & McNaughton, N. *The Neuropsychology of Anxiety: An Enquiry into the Function of the Septo-hippocampal System* Vol. 33 (Oxford University Press, 2003).
44. Frijda, N. H., Kuipers, P. & Ter Schure, E. *J. Pers. Soc. Psychol.* **57**, 212–218 (1989).
45. Grupe, D. W. & Nitschke, J. B. *Nat. Rev. Neurosci.* **14**, 488–501 (2013).
46. Kahneman, D., Slovic, P. & Tversky, A. *Judgment Under Uncertainty: Heuristics and Biases*. (Cambridge University Press, 1982).
47. FeldmanHall, O., Glimcher, P., Baker, A. L. & Phelps, E. A. *J. Exp. Psychol. Gen.* **145**, 1255–1262 (2016).
48. Bechara, A., Damasio, A. R., Damasio, H. & Anderson, S. W. *Cognition* **50**, 7–15 (1994).
49. Weary, G. & Edwards, J. A. *J. Pers. Soc. Psychol.* **67**, 308–318 (1994).
50. Kagan, J. *J. Pers. Soc. Psychol.* **22**, 51–66 (1972).
51. Thornton, M. A. & Tamir, D. I. *Proc. Natl Acad. Sci. USA* **114**, 5982–5987 (2017).
52. Cacioppo, J. T., Hughes, M. E., Waite, L. J., Hawkley, L. C. & Thisted, R. A. *Psychol. Aging* **21**, 140–151 (2006).
53. Hawkley, L. C. & Cacioppo, J. T. *Ann. Behav. Med.* **40**, 218–227 (2010).
54. Fiske, S.T. in *Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition* (eds. Clark, M. S. & Fiske, S. T.) 55–78 (Erlbaum, 1982).
55. Fiske, S.T. in *Handbook of Social Psychology* Vol. 2 (eds. Fiske, S. T., Gilbert, D. T., & Lindzey, G.) 357–411 (McGraw-Hill, 1998).
56. Fiske, S.T., Lin, M. & Neuberg, S.L. in *Dual-Process Theories in Social Psychology* (eds. Chaiken, S., & Trope, Y.) 231–254 (1999).
57. Todorov, A., Mandisodza, A. N., Goren, A. & Hall, C. C. *Science* **308**, 1623–1626 (2005).
58. Todorov, A. & Mende-Siedlecki, P. in *The Oxford Handbook of Cognitive Neuroscience* (eds. Ochsner, K. & Kosslyn, S.) Ch. 11 (Oxford University Press, 2013).
59. Tory Higgins, E., Rholes, W. S. & Jones, C. R. *J. Exp. Soc. Psychol.* **13**, 141–154 (1977).
60. Baron, S. G., Gobbini, M. I., Engell, A. D. & Todorov, A. *Soc. Cogn. Affect. Neurosci.* **6**, 572–581 (2011).
61. North, M. S., Todorov, A. & Osherson, D. N. *J. Nonverbal Behav.* **36**, 227–233 (2012).
62. Mende-Siedlecki, P., Said, C. P. & Todorov, A. *Soc. Cogn. Affect. Neurosci.* **8**, 285–299 (2013).
63. Hughes, B. L., Zaki, J. & Ambady, N. *Soc. Cogn. Affect. Neurosci.* **12**, 49–60 (2017).
64. Posner, M.I. & Snyder, C.R.R. in *Information Processing and Cognition: The Loyola Symposium* (ed. Solso, R. L.) 55–85 (Erlbaum, 1975).
65. Shiffrin, R. M. & Schneider, W. *Psychol. Rev.* **84**, 127–190 (1977).
66. Klein, N. & O'Brien, E. *Proc. Natl Acad. Sci. USA* **115**, 13222–13227 (2018).
67. van den Bos, K. *J. Pers. Soc. Psychol.* **80**, 931–941 (2001).

68. van Dijk, E., Wilke, H., Wilke, M. & Metman, L. *J. Exp. Soc. Psychol.* **35**, 109–135 (1999).
69. Bargh, J. A. & Pietromonaco, P. *J. Pers. Soc. Psychol.* **43**, 437–449 (1982).
70. Koster-Hale, J. & Saxe, R. *Neuron* **79**, 836–848 (2013).
71. Schaafsma, S. M., Pfaff, D. W., Spunt, R. P. & Adolphs, R. *Trends Cogn. Sci.* **19**, 65–72 (2015).
72. Baker, C.L., Saxe, R. & Tenenbaum, J.B. Bayesian theory of mind: modelling joint belief-desire attribution. *Proceedings of the Annual Meeting of the Cognitive Science Society* (2011).
73. Devaine, M., Hollard, G. & Daunizeau, J. *PLOS Comput. Biol.* **10**, e1003992 (2014).
74. Devaine, M. et al. *PLOS Comput. Biol.* **13**, e1005833 (2017).
75. Zaki, J. & Ochsner, K. N. *Nat. Neurosci.* **15**, 675–680 (2012).
76. FeldmanHall, O., Dalgleish, T., Evans, D. & Mobbs, D. *Neuroimage* **105**, 347–356 (2015).
77. Olsson, A. et al. *Psychol. Sci.* **27**, 25–33 (2016).
78. Miller, S. A. *Psychol. Bull.* **135**, 749–773 (2009).
79. Wellman, H. M., Cross, D. & Watson, J. *Child Dev.* **72**, 655–684 (2001).
80. Fiske, S. T. & Taylor, S.E. *Social Cognition* (McGraw-Hill, 1991).
81. Zaki, J. *Trends Cogn. Sci.* **21**, 59–60 (2017).
82. Cameron, D., Hutcherson, C., Ferguson, A., Scheffer, J. & Inzlicht, M. Empathy is a choice: people are empathy misers because they are cognitive misers. *Rotman School of Management Working Paper No. 2887903* (2016).
83. Cohen, J. D., Dunbar, K. & McClelland, J. L. *Psychol. Rev.* **97**, 332–361 (1990).
84. MacLeod, C. M. & Dunbar, K. *J Exp. Psychol. Learn.* **14**, 126–135 (1988).
85. Shenhav, A. *Psychol. Inq.* **28**, 148–152 (2017).
86. Mata, R., Josef, A. K., Samanez-Larkin, G. R. & Hertwig, R. *Ann. NY Acad. Sci.* **1235**, 18–29 (2011).
87. Grubb, M. A., Tymula, A., Gilaie-Dotan, S., Glimcher, P. W. & Levy, I. *Nat. Commun.* **7**, 13822 (2016).
88. Tymula, A., Rosenberg Belmaker, L. A., Ruderman, L., Glimcher, P. W. & Levy, I. *Proc. Natl Acad. Sci. USA* **110**, 17143–17148 (2013).
89. Surtees, A., Apperly, I. & Samson, D. *Cognition* **150**, 43–52 (2016).
90. Gallagher, H. L. & Frith, C. D. *Trends Cogn. Sci.* **7**, 77–83 (2003).
91. Singer, T. *Neurosci. Biobehav. Rev.* **30**, 855–863 (2006).
92. Singer, T. et al. *Science* **303**, 1157–1162 (2004).
93. Neumann, R. & Strack, F. *J. Pers. Soc. Psychol.* **79**, 211–223 (2000).
94. Barsade, S. G. *Adm. Sci. Q.* **47**, 644–675 (2002).
95. Frith, C. D. & Frith, U. *Brain Res.* **1079**, 36–46 (2006).
96. Gintis, H. & Fehr, E. *Behav. Brain Sci.* **35**, 28–29 (2012).
97. Bandura, A. *Adv. Exp. Soc. Psychol.* **2**, 1–55 (1965).
98. Behrens, T. E., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. *Nature* **456**, 245–249 (2008).
99. Debiec, J. & Olsson, A. *Trends Cogn. Sci.* **21**, 546–555 (2017).
100. Apps, M. A., Lesage, E. & Ramnani, N. *J. Neurosci.* **35**, 2904–2913 (2015).
101. FeldmanHall, O. & Dunsmoor, J. E. *Perspect. Psychol. Sci.* **14**, 175–196 (2018).
102. FeldmanHall, O., Dunsmoor, J. E., Kroes, M. C. W., Lackovic, S. & Phelps, E. A. *Psychol. Sci.* **28**, 1160–1170 (2017).
103. Nassar, M. R., Wilson, R. C., Heasly, B. & Gold, J. I. *J. Neurosci.* **30**, 12366–12378 (2010).
104. Courville, A. C., Daw, N. D. & Touretzky, D. S. *Trends Cogn. Sci.* **10**, 294–300 (2006).
105. Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E. & Cohen, J. D. *Science* **300**, 1755–1758 (2003).
106. Tomlin, D. et al. *Science* **312**, 1047–1050 (2006).
107. Mobbs, D. et al. *Soc. Cogn. Affect. Neurosci.* **10**, 1323–1328 (2015).
108. Apps, M. A., Rushworth, M. F. & Chang, S. W. *Neuron* **90**, 692–707 (2016).
109. Apps, M. A. J. & Sallet. *J. Trends Cogn. Sci.* **21**, 151–152 (2017).
110. Sutton, R. S. B. *Reinforcement Learning: An Introduction*. (MIT Press, 1998).
111. Laquitaine, S. & Gardner, J. L. *Neuron* **97**, 462–474.e6 (2018).
112. Lieder, F., Griffiths, T. L., M Huys, Q. J. & Goodman, N. D. *Psychon. Bull. Rev.* **25**, 322–349 (2018).
113. Bogacz, R. *Trends Cogn. Sci.* **11**, 118–125 (2007).
114. Ratcliff, R., Smith, P. L., Brown, S. D. & McKoon, G. *Trends Cogn. Sci.* **20**, 260–281 (2016).
115. Hogg, M. A. *Eur. Rev. Soc. Psychol.* **11**, 223–255 (2000).
116. Whalen, P. J. *Curr. Dir. Psychol. Sci.* **7**, 177–188 (1998).
117. Fischbacher, U., Gächter, S. & Fehr, E. *Econ. Lett.* **71**, 397–404 (2001).
118. Mullin, B. A. & Hogg, M. A. *Basic Appl. Soc. Psych.* **21**, 91–102 (1999).
119. Hogg, M. A. *Eur. Rev. Soc. Psychol.* **4**, 85–111 (1993).
120. Epley, N. & Gilovich, T. *Psychol. Sci.* **12**, 391–396 (2001).
121. Tversky, A. & Kahneman, D. *Science* **185**, 1124–1131 (1974).
122. van den Berg, P. & Wenseleers, T. *Nat. Commun.* **9**, 2151 (2018).
123. Tamir, D. I. & Mitchell, J. P. *J. Exp. Psychol. Gen.* **142**, 151–162 (2013).
124. Lieder, F., Shenhav, A., Musslick, S. & Griffiths, T. L. *PLOS Comput. Biol.* **14**, e1006043 (2018).
125. Shenhav, A. et al. *Annu. Rev. Neurosci.* **40**, 99–124 (2017).
126. Epley, N. & Gilovich, T. *Psychol. Sci.* **17**, 311–318 (2006).
127. Mendes, W. B., Blascovich, J., Hunter, S. B., Lickel, B. & Jost, J. T. *J. Pers. Soc. Psychol.* **92**, 698–716 (2007).
128. Suzuki, S., Jensen, E. L. S., Bossaerts, P. & O'Doherty, J. P. *Proc. Natl Acad. Sci. USA* **114**, E5278–E5278 (2017).
129. Kruglanski, A. W., Pierro, A., Mannetti, L. & De Grada, E. *Psychol. Rev.* **113**, 84–100 (2006).
130. Hogg, M. A., Sherman, D. K., Dierselhuis, J., Maitner, A. T. & Moffitt, G. *J. Exp. Soc. Psychol.* **43**, 135–142 (2007).
131. Vives, M. L. & FeldmanHall, O. *Nat. Commun.* **9**, 2156 (2018).
132. Hogg, M. A., Adelman, J. R. & Blagg, R. D. *Pers. Soc. Psychol. Rev.* **14**, 72–83 (2010).
133. de Coppet, D. *Understanding Rituals*. (Routledge, 2002).
134. Hobson, N. M., Schroeder, J., Risen, J. L., Xygalatas, D. & Inzlicht, M. *Pers. Soc. Psychol. Rev.* **22**, 260–284 (2018).
135. Brooks, A. W. et al. *Organ. Behav. Hum. Decis. Process* **137**, 71–85 (2016).
136. Norton, M. I. & Gino, F. *J. Exp. Psychol. Gen.* **143**, 266–272 (2014).
137. Finn, E. S., Corlett, P. R., Chen, G., Bandettini, P. A. & Constable, R. T. *Nat. Commun.* **9**, 2043 (2018).
138. Delton, A. W., Krasnow, M. M., Cosmides, L. & Tooby, J. *Proc. Natl Acad. Sci. USA* **108**, 13335–13340 (2011).
139. McNamara, J. M., Barta, Z. & Houston, A. I. *Nature* **428**, 745–748 (2004).
140. Vaghi, M. M. et al. *Neuron* **96**, 348–354.e4 (2017).
141. Carleton, R. N. *J. Anxiety Disord.* **41**, 5–21 (2016).
142. Carleton, R. N. *Expert Rev. Neurother.* **12**, 937–947 (2012).
143. Engelmann, J. B., Meyer, F., Fehr, E. & Ruff, C. C. *J. Neurosci.* **35**, 3085–3099 (2015).
144. Carleton, R. N., Collimore, K. C. & Asmundson, G. J. G. *J. Anxiety Disord.* **24**, 189–195 (2010).
145. Yoon, K. L. & Zinbarg, R. E. *J. Abnorm. Psychol.* **117**, 680–685 (2008).
146. Blanchette, I. & Richards, A. *J. Exp. Psychol. Gen.* **132**, 294–309 (2003).
147. Mogg, K., Bradbury, K. E. & Bradley, B. P. *Behav. Res. Ther.* **44**, 1411–1419 (2006).
148. Gentes, E. L. & Ruscio, A. M. *Clin. Psychol. Rev.* **31**, 923–933 (2011).
149. Norton, R. W. *J. Pers. Assess.* **39**, 607–619 (1975).
150. Roemer, L. & Orsillo, S. M. *Clin. Psychol. Sci. Pract.* **9**, 54–68 (2002).
151. Boucher, J. *Br. J. Disord. Commun.* **24**, 181–198 (1989).
152. Wigham, S., Rodgers, J., South, M., McConachie, H. & Freeston, M. *J. Autism Dev. Disord.* **45**, 943–952 (2015).
153. Sevgi, M., Diaconescu, A. O., Tittgemeyer, M. & Schilbach, L. *Biol. Psychiatry* **80**, 112–119 (2016).
154. Eil, D. & Rao, J. M. *Am. Econ. J. Microecon.* **3**, 114–138 (2011).
155. Persoskie, A., Ferrer, R. A. & Klein, W. M. P. *J. Behav. Med.* **37**, 977–987 (2014).
156. Ganguly, A. & Tasoff, J. *Manage. Sci.* **63**, 4037–4060 (2017).
157. Charpentier, C. J., Bromberg-Martin, E. S. & Sharot, T. *Proc. Natl Acad. Sci. USA* **115**, E7255–E7264 (2018).
158. Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. *Nature* **441**, 876–879 (2006).
159. Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A. & Cohen, J. D. *J. Exp. Psychol. Gen.* **143**, 2074–2081 (2014).
160. Yamagishi, T., Cook, K. S. & Watabe, M. *Am. J. Sociol.* **104**, 165–194 (1998).
161. Wittmann, B. C., Daw, N. D., Seymour, B. & Dolan, R. J. *Neuron* **58**, 967–973 (2008).
162. Kakade, S. & Dayan, P. *Neural Netw.* **15**, 549–559 (2002).
163. Cohen, J. D., McClure, S. M. & Yu, A. J. *Phil. Trans. R. Soc. Lond. B* **362**, 933–942 (2007).
164. Schmidhuber, J. in *Proceedings of the International Conference on Simulation of Adaptive Behavior: From Animals to Animats* (eds. Meyer, J. A. & Wilson, S. W.) 222–227 (MIT Press/Bradford Books, 1991).
165. Eastwood, J. D., Frischen, A., Fenske, M. J. & Smilek, D. *Perspect. Psychol. Sci.* **7**, 482–495 (2012).
166. Schwartenbeck, P., Fitzgerald, T., Dolan, R. J. & Friston, K. *Front. Psychol.* **4**, 710 (2013).
167. Jepma, M., Verdonschot, R. G., van Steenbergen, H., Rombouts, S. A. & Nieuwenhuis, S. *Front. Behav. Neurosci.* **6**, 5 (2012).
168. Loewenstein, G. F. *Psychol. Bull.* **116**, 75–98 (1994).
169. Berlyne, D. E. *J. Exp. Psychol.* **53**, 399–404 (1957).
170. Kruglanski, A. W. *Psychol. Inq.* **1**, 181–197 (1990).
171. Kruglanski, A. W. & Freund, T. *J. Exp. Soc. Psychol.* **19**, 448–468 (1983).
172. Kruglanski, A. W., Peri, N. & Zakai, D. *Soc. Cogn.* **9**, 127–148 (1991).
173. Dijksterhuis, A., vanKnippenberg, A., Kruglanski, A. W. & Schaper, C. *J. Exp. Soc. Psychol.* **32**, 254–270 (1996).
174. Richter, L. & Kruglanski, A. W. *J. Exp. Soc. Psychol.* **34**, 313–329 (1998).
175. Festinger, L. *A Theory of Cognitive Dissonance*. (Stanford University Press, 1957).
176. Kruglanski, A. W., Shah, J. Y., Pierro, A. & Mannetti, L. *J. Pers. Soc. Psychol.* **83**, 648–662 (2002).
177. Hogg, M. A. *Eur. Psychol.* **9**, 284–285 (2004).

178. Heine, S. J., Proulx, T. & Vohs, K. D. *Pers. Soc. Psychol. Rev.* **10**, 88–110 (2006).
179. Mitchell, J. P. *Trends Cogn. Sci.* **13**, 246–251 (2009).
180. Tamir, D. I. & Thornton, M. A. *Trends Cogn. Sci.* **22**, 201–212 (2018).
181. Barrett, L. F. *Soc. Cogn. Affect. Neurosci.* **12**, 1833 (2017).
182. Baker, C. L., Saxe, R. & Tenenbaum, J. B. *Cognition* **113**, 329–349 (2009).
183. Berlyne, D. E. *Psychol. Rev.* **64**, 329–339 (1957).
184. Keramati, M., Dezfouli, A. & Piray, P. *PLOS Comput. Biol.* **7**, e1002055 (2011).
185. Miller K., Shenhav, A. & Ludvig, E. Habits without values. Preprint at *bioRxiv* https://www.biorxiv.org/content/10.1101/067603v4 (2017).
186. Graybiel, A. M. *Annu. Rev. Neurosci.* **31**, 359–387 (2008).
187. Gillan, C. M. et al. *Am. J. Psychiatry* **168**, 718–726 (2011).
188. Lawson, R. P., Rees, G. & Friston, K. J. *Front. Hum. Neurosci.* **8**, 302 (2014).
189. Friston, K. *Biol. Psychiatry* **80**, 87–89 (2016).
190. Cimino, A. *J. Cogn. Cult.* **11**, 241–267 (2011).
191. Dickinson, A. *Phil. Trans. R. Soc. Lond. B* **308**, 67–78 (1985).
192. Griffiths, T. L. & Tenenbaum, J. B. *Psychol. Sci.* **17**, 767–773 (2006).
193. Pezzulo, G. *Cogn. Affect. Behav. Neurosci.* **14**, 902–911 (2014).
194. Kahneman, D. *Am. Psychol.* **58**, 697–720 (2003).
195. Epley, N. & Gilovich, T. *Pers. Soc. Psychol. Bull.* **30**, 447–460 (2004).
196. Epley, N., Keysar, B., Van Boven, L. & Gilovich, T. *J. Pers. Soc. Psychol.* **87**, 327–339 (2004).
197. Epley, N. & Dunning, D. *Pers. Soc. Psychol. Bull.* **32**, 641–655 (2006).
198. Stern, C. & West, T. V. *Pers. Soc. Psychol. Bull.* **42**, 1466–1479 (2016).
199. Kruger, J. & Dunning, D. *J. Pers. Soc. Psychol.* **77**, 1121–1134 (1999).
200. Sporer, S. L., Penrod, S., Read, D. & Cutler, B. *Psychol. Bull.* **118**, 315–327 (1995).

Acknowledgements

We thank C. Dean Wolf for creation of the figures and thank M. Frank and D. Tamir for helpful comments on early drafts. This research was supported by a Center of Biomedical Research Excellence grant (P20GM103645) from the National Institute of General Medical Sciences to O.F.H. and A.S.

Competing interests

The authors declare no competing interests.

Additional information

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence should be addressed to O.F.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.