The Organization of Social Knowledge Is Tuned for Prediction

Mark A. Thornton and Diana I. Tamir

The social world requires people to make highly consequential predictions. People need to predict whether new acquaintances will become a friend or foe (Cuddy, Fiske, & Glick, 2008), how old friends will respond to constructive criticism, or how long the boss will be angry before approaching her for a favor. Whether in cooperation or competition, successful social interaction requires people to anticipate others' future thoughts, feelings, and actions, and prepare their own actions accordingly. Social predictions are among the most common predictions a person must make because people spend so much time with other people (United States Bureau of Labor Statistics, 2003). Yet despite the importance of social prediction, researchers have only just scratched the surface of the predictive social mind. Here we consider recent research that is starting to reveal how people glimpse the social future.

Research on nonsocial prediction suggests that the brain is built to make predictions. It does not passively perceive the world around it and then react accordingly. Instead, people make reflexive predictions across multiple domains (Hohwy, Roepstorff, & Friston, 2008; Rao & Ballard, 1999; Vuust, Ostergaard, Pallesen, Bailey, & Roepstorff, 2009). When processing language, for example, people use the beginning of a sentence to predict the end of that sandwich. When watching a ball thrown into the air, people reflexively make a prediction about its eventual downward trajectory. However, the social world poses unique challenges to people's well-honed predictive capacities. Humans are not billiard balls: people are probabilistic beings, moved to action by the unseen forces of thoughts and feelings.

M. A. Thornton (✉)
Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA

D. I. Tamir
Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA
e-mail: mark.a.thornton@dartmouth.edu; dtamir@princeton.edu

M. Gilead, K. N. Ochsner (eds.), *The Neural Basis of Mentalizing*,
https://doi.org/10.1007/978-3-030-51890-5_14

How do people represent these invisible mental states, and use them to make predictions?

A recent theoretical framework for social cognition (Tamir & Thornton, 2018) proposes a simple answer to this question (Fig. 1). This multilayered framework of

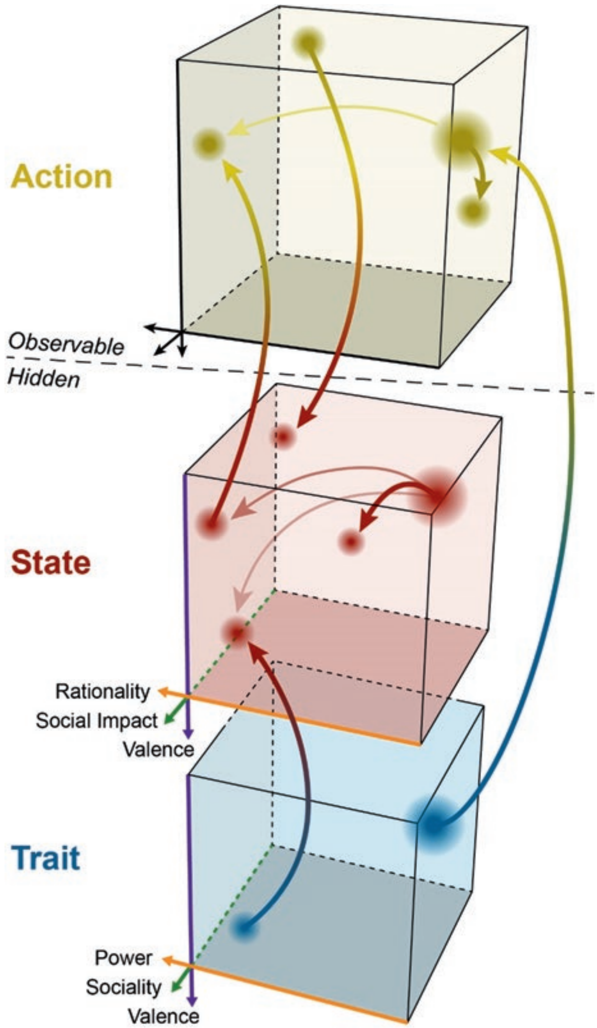

Fig. 1 The multilayer model of predictive social cognition. Three layers of social knowledge—observable actions and hidden mental states and traits—are each organized into low-dimensional maps by psychological dimensions such as valence and power. Transitions between or within layers (arrows) decrease in probability with distance. Short hops between adjacent points (e.g., happiness and gratitude) are more likely than long treks between distant points (e.g., sleeping and hiking). This organization of social knowledge provides both for parsimonious representation of the complexities of the social world, and for accurate, automatic prediction of the social future. (Reproduced with permission from *Trends in Cognitive Science*)

social cognition helps to explain how people predict others' future states and behaviors in two steps: First, it suggests that the mind organizes social knowledge using conceptual "maps" of social stimuli. These maps allow people to easily track other people's current thoughts, feeling, and actions. Second, it suggests that people track distances and trajectories through these maps to make efficient, automatic social predictions. This framework advances prediction as the central goal of representing social knowledge.

In this chapter, we first describe how people simplify the complexity of the social world using these low-dimensional maps. Next, we discuss how people leverage these maps to make accurate, automatic social predictions. Finally, we offer suggestions for how future research can use this framework productively to model real-world social predictions, and constructively to enhance its explanatory power and scope.

The Organization of Social Knowledge

Humans enjoy extraordinarily rich social lives. Most people know hundreds or even thousands of unique individuals. In a given moment, each one of these individuals may be thinking widely differing thoughts, experiencing different emotions, or performing different actions. This richness makes life interesting and exciting, but it also poses a significant challenge: people must understand others' traits, mental states, and action in order to successfully interact with them. How can the social mind come to grips with the complexity of social life?

The right organizational scheme can bring order to chaos. What kind of scheme does the human mind use to organize social information? We propose that people use a social map. Or rather, that people employ multiple maps, one for each type of social information. Maps organize information by localizing it to particular coordinates on a small set of continuous dimensions. In a geographic map of the United States, the physical location of any city can be described in terms of its dimensions of longitude and latitude. Simply knowing the north-south and east-west coordinates of a city allows you to extract important information about its location from single pair of numbers. For instance, you might learn that a city is located in the Pacific Northwest, and thus that it is nearby to Seattle, WA, but far away from Miami, FL.

Conceptual maps of the social world act in a similar way, by reducing the complexity of social stimuli down to a few essential values. Social neuroscience research has begun to chart the maps that the brain uses to make sense of the social world. This work has revealed the cardinal dimensions that describe three key types of social information: actions, mental states, and traits. These three layers of social information form the core of what one might want to know about a person: what are they doing, how are they feeling, and what kind of person are they? Each layer captures the world at a different timescale. Actions occur at the shortest timescales, from less than a second up to a few hours, depending on their complexity. States might unfold over just a few minutes or persist for several days, depending on if they are more like emotions or moods. Traits are more lasting, or even permanent, ways to describe individual differences among people. To understand and predict other people at any time scale requires us to map out the content of each of these layers of social knowledge.

Mapping the Action Layer

Humans have an incredibly diverse behavioral repertoire. People are capable of engaging in thousands of different actions and activities, ranging from simple motor actions such as reaching and grasping, to complex extended activities such as conducting research or governing a nation. A successful social agent must have a keen understanding of these actions. Recent research (Thornton & Tamir, 2019b) has identified six psychological dimensions (Table 1) that scaffold people's action concepts: the Abstraction, Creation, Tradition, Food, Animacy, and Spiritualism Taxonomy or ACT-FAST. These dimensions originated from a data-driven principal component analysis of verb use in large text corpora and were validated against behavioral judgments.

Like the other social maps we will explore, each dimension of the ACT-FAST carries intrinsic meaning. For example, if one knows that an action is high on the animacy dimension, then one knows that it is an action that tends to be performed by living agents such as humans and animals, rather than an act of nature or a machine. In this way, this conceptual map of the social world both offers the functionality of a physical map—representing the distances between difference locations—and also implies rich knowledge about each of those locations.

Together, the six ACT-FAST dimensions explain much of how people think about actions. Knowing an action's coordinates on these dimensions can robustly predict: (1) *who* does an action, in terms of traits, (2) *why* one does an action, in terms of approach and avoidance motivations, (3) *when* one does an action, in terms of time of day, (4) *where* one does an action, in terms of outdoor versus indoor, and if indoor, public versus private, and (5) *how* one performs an action, in terms of body parts involved and mental or physical effort required. ACT-FAST can also explain patterns of natural language use, such as which verbs tend to co-occur with

Table 1 Dimensions of the FAACTS action taxonomy

Dimension	Pole 1	Pole 2	Examples
Food	Food	Nonfood	Bake, fry vs. detain, testify
Abstraction	Abstract/social	Concrete/physical	Govern, refute vs. drip, peel
Animacy	Animate	Mechanical	Meow, floss vs. contain, extract
Creation	Creation	Crime	Film, sing vs. prosecute, testify
Tradition	Tradition	Innovation	Cook, decorate vs. emit, encrypt
Spiritualism	Work	Worship	Fax, haggle vs. foretell, ascend

which nouns. In addition to answering these important psychological and linguistic questions, ACT-FAST can explain brain activity across a wide set of cortical regions implicated in action representation. Together, these findings suggest that ACT-FAST provide a useful, and biologically plausible map of how people organize knowledge about other people's actions.

Mapping the State Layer

In addition to paying attention to others' observable actions, successful social agents also attend to the hidden drivers of actions: mental states. These states must be inferred from indirect cues, such as facial expression and tone of voice, but once known, others' thoughts and feelings can serve as powerful predictors of their behavior. People often share the same intuitions about the predictive power of mental states: angry people aggress, tired people rest, and happy people celebrate. Failing to attend to others' mental states could lead to embarrassing faux pas at best, or to serious danger at worst. To avoid such pitfalls, we need a map of the state layer.

A number of established theories propose dimensions that organize mental states. For example, the circumplex model of affect offers a map with two dimensions, valence and arousal (Russell, 1980); or the distinction between emotional and rational states, prominent in many modern dual-process theories (Evans & Stanovich, 2013; Kahneman, 2003). Recent neuroimaging work (Tamir, Thornton, Contreras, & Mitchell, 2016) used PCA to synthesize the 16 dimensions that comprised these existing theories, and then validated the resulting components in terms of their ability to predict neural representations of mental states. This work identified three dimensions that structure the brain's map of others' mental states (Table 2). The first dimension, valence, captures whether others are feeling good or bad. Knowing the valence of a person's mental state could help people avoid harm from those in negative states like rage, and enjoy pleasant, constructive social interactions with those in more positive moods. The second dimension on this map, social impact, captures which mental states would dispose others to engage in social interactions. Highly impactful states, whether good or bad, are more likely to affect one's life. The final dimension, rationality, captures whether others are likely to act in a calm, deliberate, well-thought-out way, or react instinctively or rashly.

Valence, social impact, and rationality together comprise the 3d Mind Model. This model can explain over 80% of the variance in neural representations of mental states. That is, it provides a near-complete map of the mental state layer (Thornton

Table 2 Dimensions of the mental state representation

Dimension	Pole 1	Pole 2	Examples
Valence	Positive	Negative	Ecstasy, peacefulness vs. rage, sadness
Social impact	Impactful	Unimpactful	Envy, love vs. exhaustion, self-pity
Rationality	Rational	Emotional	Thinking, planning vs. joy, lust

& Tamir, 2020a). Moreover, this map remains robust across different ways of perceiving mental states: similar dimensions emerge across modalities, regardless of whether people reflect on mental state-related scenarios presented as images or as text (Weaverdyck, Thornton, & Tamir, 2020). People also apply similar maps to thinking about their own minds, as opposed to the minds of others. While the structure of the map doesn't change across targets, the resolution does: when people think about their own mental states, they pore over a highly detailed, richly annotated map (Thornton, Weaverdyck, Mildner, & Tamir, 2019). In contrast, when people think about the states of others, the resolution is much less fine-grained. This difference likely results from both the quality and quantity of information one has about their own minds, in contrast to the minds of others.

Mapping the Trait Layer

Although people may apply similar maps to the mental states of different people, individuals do differ in socially important ways. Enduring individual differences between people are known as traits. Compared with actions and mental states, traits are relatively permanent fixtures of an individual, changing slowly across a lifetime, if at all (Roberts & Mroczek, 2008; Srivastava, John, Gosling, & Potter, 2003). Knowing where a person places on trait dimensions can help people to make predictions about their likely states or actions. For example, if you know that someone is highly trustworthy, you can predict that they will not steal from you; if you know someone is highly social, you can predict that they might feel excited at a party. Traits thus help people to make nuanced predictions about people, across situations.

There are multiple existing dimensional maps of traits, including the Five Factor model (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) of personality (Goldberg, 1990; McCrae & Costa, 1987), and the stereotype content model (Fiske, Cuddy, Glick, & Xu, 2002) consisting of warmth and competence. Recent neuroimaging research synthesized several of the most prominent trait theories in the literature (Thornton & Mitchell, 2018). In this work, a three-dimensional model—consisting of power, valence, and sociality—provided the best explanation for patterns of brain activity elicited by thinking about a large set of public figures. Knowing where a person resides on these dimensions can inform social judgments about them. Power indicates whether another person is dominant and competent, and thus, capable of enacting their will. Valence indicates whether that person is warm and trustworthy, and thus, likely to help or harm. Finally, sociality reflects whether a person is extraverted, and thus, likely to engage in the first place. Together, these dimensions provide a near-complete map of the trait space: Power, valence, and sociality explain more than two-thirds of reliable neural activity associated with making inferences about other people. This model outperforms even the Big 5 personality traits in explaining neural representations of other people. However, it is important to note that the Big 5 are ostensibly a model of the reality of traits, whereas the three-dimensional model instead aims to explain the perception of traits. Moreover, it is an open question whether people continue to apply any map of trait space for personally familiar others (Thornton & Mitchell, 2017). When people interact regularly, they may instead draw on the character of their relationship itself.

Overlapping Social Maps

The maps our mind makes of the social world rely on a similar set of brain regions. Both mental state and trait representation engage regions such as the medial prefrontal and parietal cortices, and the superior temporal sulcus extending from the temporoparietal junction forward to the anterior temporal lobe (Tamir et al., 2016; Thornton & Mitchell, 2018). These regions form the social brain network, a set of brain regions reliably activated by a wide range of social stimuli (Mitchell, 2008; Van Overwalle & Baetens, 2009). However, mental states and traits share more than just a gross anatomical similarity—they share a common neural code. Both maps include a dimension reflecting valence; both include a dimension reflecting sociality; and both include a dimension reflecting competence, dominance, and agency. This allows each dimension to be decoded from brain activity across domains (Thornton & Mitchell, 2018). This shared code hints at a deep connection between the way people think about others' momentary and enduring mental properties—potentially bridging the traditional divide between traits and states.

Indeed, in recent research, we found that when our brain represents a person, it seems to do so by keeping track of the mental states that person habitually experiences (Thornton, Weaverdyck, & Tamir, 2019a). For example, if a politician is habitually in bad moods—grouchy, short-tempered, stubborn—one may form the impression that the politician has a negative disposition, at the trait level. Correspondingly, our results indicate that the pattern elicited by thinking about such a person could be reconstructed by adding together generic representations of grouchiness, short-temper, and stubbornness. This finding suggests a simple mechanism for impression formation—counting perceived mental states. Moreover, this process could later be reversed to make predictions about states based on people's dispositions. Thus, although traits and states are typically thought of as separate, the trait and state dimensions described above reflect parallel concepts.

However, not all socially relevant information occupies the same neural territory. In recent research, we and others have found that actions are represented in quite different portions of the brain than those involved in theory of mind (Tarhan & Konkle, 2019; Thornton & Tamir, 2020b). High-level visual regions within the dorsal and ventral paths—areas rarely implicated in social cognition per se—appear to play an important role in representing what others are doing. Understanding more about how maps of the social world relate—both in conceptual space and the physical territory of the brain—is a high priority for future research.

The Predictive Social Brain

Conceptual maps of actions, mental states, and traits allow people to organize their knowledge of the social world. However, the social world is not static; it is constantly in flux. At one moment, a person may be having fun dancing with their friends, but after a few hours of this energetic state, they may find themselves feeling exhausted. This mental state of exhaustion may in turn lead to a new activity, like resting, and so forth. Navigating the social world requires people to anticipate such transitions. Fortunately, the dimensional maps of social knowledge described above offer insight into these dynamics: knowing "where" a person is on a social map can tell you a great deal about where that person will "go" in the future.

Social Dimensions Scaffold Social Predictions

How can the static maps described above be used to predict the social future? The key assumption that licenses such predictions is that the proximity between points on the map reflects the likelihood of moving between those points. This assumption holds for geographic maps: on average, people tend to travel locally—e.g., to the store, work, school, or gym—with relatively high frequency. People travel to more distant locales—another city, country, or continent—much more rarely. Consequently, if you know another person's current location on a map, you can accurately predict that their next destination is going to be close by.

We propose that people apply an analogous algorithm to make social predictions using conceptual maps. The coordinates on these maps represent actions, mental states, and personality traits rather than physical locations, but the logic of distance-based prediction still applies. Simply put, people are more likely to transition from one location within a layer to a nearby location than they are to transition to faraway locations. For example, if a person is currently feeling pride, a positive emotion, one might guess that she is much more likely to feel happy next than to feel sad. In this way, one could predict the likely trajectory of a person through a layer simply by knowing that person's current coordinates. If so, then the multilayer structure mapped above would provide the foundation for understanding the dynamics of the social world.

Behavioral research suggests that people do predict the likelihood of transitions between states based on the proximity between those states on the dimensions of the 3d Mind Model—valence, social impact, and rationality. To demonstrate this, we first elicited people's intuitions for how states transition from one to the next (Thornton & Tamir, 2017). For example, a participant might be told that a person is currently feeling "excitement" and be asked to rate the likelihood that that person will next experience "sleepiness" from 0% to 100%. Each state was mapped according to its location on the dimensions of rationality, social impact, and valence so that we could calculate the "distance" between each pair of states. Across two studies,

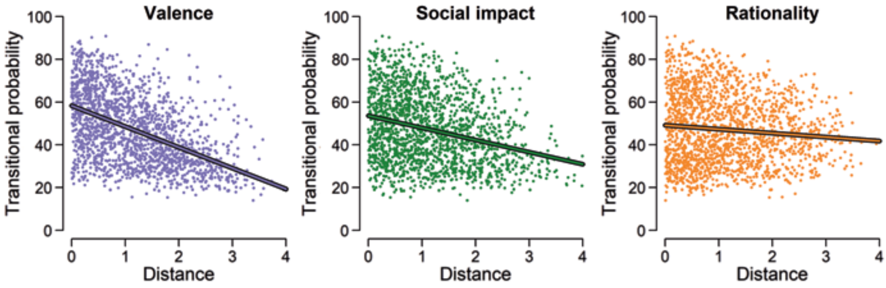

Fig. 2 Dimensional proximity predicts transitional probabilities between states. Each point on the scatter plots represents a transition from one mental state to another. The *x*-axis indicates the absolute distance between those states on each mental state dimension. The *y*-axis indicates the predicted transitional probability from one state to the other. The further two states are on each dimension, the less people expect a transition from one to the next. People are also less likely to actually transition between distant states. These state dimensions explain much of the accuracy of predicted transitional probabilities (Thornton & Tamir, 2017)

proximity on the dimensions was positively associated with participants' transitional probability ratings. The closer states were to each other on any of these mental state dimensions, the more likely participants judged the transitions between them (Fig. 2).

In the same way, people use conceptual proximity on the ACT-FAST dimensions to predict others' actions (Thornton & Tamir, 2019a). Across five preregistered behavioral experiments, participants rated the likelihood that a person currently engaged in one action to next engage in another. For instance, how likely would it be for someone currently "dancing" to next "rest"? Proximities between actions on the ACT-FAST action dimensions reliably predict people's transitional probability ratings. As with states, this result suggests that people may draw upon their map of the action layer to make predictions about others' likely future actions, based on their current actions.

Social Dimensions Describe Real Social Dynamics

In the previous section, we described evidence that people use proximity within a social map to make predictions about others' mental states and actions. However, it would only make sense to use social maps to make predictions if the social dimensions describe actual social dynamics. That is, one should use proximity on the dimension of valence to predict mental state transitions if and only if valence in fact describes regularities in the mental state transitions that others actually experience. As part of the investigations described in the previous section, we used experience sampling data and other real-world data to measure actual state and action transitions. This allowed for an estimation of the actual transitional probabilities between pairs of mental states (Thornton & Tamir, 2017) and pairs of actions (Thornton & Tamir, 2019a).

Distance on the mental state map predicted actual mental state transitions. The further away two states were on any dimension, the less likely people were to transition between them. Thus, these dimensions likely serve as a scaffolding for social prediction because they describe experienced social dynamics. Not only were perceived and actual emotion transitions correlated, but these associations could be partially explained by how close the emotions were on the dimensions of mental state representation described above. This indicates that people use their maps of mental state space to accurately predict others' emotion dynamics.

Subsequent research found similar results in the action layer (Thornton & Tamir, 2019a). Across five studies, distance on the action map predicted actual action transitions. The further two actions were on the ACT-FAST dimensions, the less likely people were to transition between them. Moreover, as in the case of mental state representation, people were highly accurate about actual action dynamics, and distance within the action map statistically mediated much of the association between perceived and actual action transitions. This finding adds further weight to the contention that people use their maps of the social world to make accurate social predictions.

The ground truth in all the studies above reflects transitional probabilities aggregated across many people or datasets. These data demonstrate that people make accurate predictions about a "generic" other. More recent data suggest that people can likewise make accurate predictions about both specific people and relevant social groups (Zhao, Thornton, & Tamir, 2018). For example, in one set of studies, we asked undergraduates to make predictions about a specific other—either a close friend, or their current roommate—as well as their undergraduate community in general. In all cases, people were able to accurately and specifically predict their friend, their roommate, and their community. This indicates that people may have highly accurate models of state dynamics in general, and that they can also tailor these models to make predictions about the individuals in their lives.

People Make Social Predictions Automatically

People can predict the social future with such high fidelity because prediction is built into the way that people represent social knowledge. In recent research from the action domain, we found that while people watch a movie, their brains spontaneously encode the actions they perceive on the ACT-FAST dimensions, and these ACT-FAST coordinates predict actions later in the movie (Thornton & Tamir, 2020b). That is, a participant's brain activity at a given moment in time automatically predicts the actions which are actually likely to occur later on. For example, if participants saw a person "running," then they would encode this action on ACT-FAST—attributing to this act a high degree of animacy. This would, in turn, accurately predict that they were likely to see other high animacy actions in the near future. Merely by encoding perceived actions on an appropriate set of dimensions—dimensions which through real actions flow smoothly—the brain thus automatically predicts likely future actions.

Similar evidence for the automaticity of social prediction comes from the domain of mental states. Whenever someone thinks about a mental state, they do not think about that mental state in isolation; they also spontaneously think about likely future states. For example, when one observes a friend experiencing pride, they can accurately predict that the friend will soon feel happy because the representation of pride incorporates the representation of happiness. Neuroimaging work has provided a unique source of evidence that this is the case: neural patterns associated with a mental state currently under consideration literally resembled patterns of likely future states (Thornton, Weaverdyck, & Tamir, 2019b). The more likely one state is to transition to another state, the more similar the neural patterns that represent them. Importantly, this work also showed that transition predictions, and not simply similarity, drove this neural finding. Even though similarity and transition likelihood are highly intertwined concepts, multiple lines of evidence suggest that transitions, and not similarity, may be primary in defining the conceptual space of mental states.

Indirect neural evidence also supports the automaticity of social prediction. In the same study, repetition suppression (also known as fMRI adaptation) tested how the brain reacted to expected and unexpected sequences of mental states. The principle behind this analysis is that, if the brain is constantly making automatic predictions, perceiving information that violates these predictions should elicit more activity than perceiving prediction-consistent information, since the latter requires recalibration of subsequent predictions. In line with this hypothesis, the study found that seeing states in unexpected sequences elicited more activity in the precuneus than observing predictable state sequences. This finding suggests that this brain region might automatically track errors in mental state predictions and update subsequent predictions accordingly.

People's maps of the social world play a key role in making the brain's social predictions automatic. Since the proximity on dimensions such as rationality, social impact, and valence is associated with transitional probabilities, as described in the previous sections, then merely encoding a state using these dimensions implicitly makes a prediction. That is, simply specifying the location of a state on such dimensions provides an indication of which other states are more or less likely. Supporting this idea, this study also found that proximity on these three dimensions of mental state representation statistically mediated much of the relationship between transitional probabilities and neural pattern similarity (Thornton, Weaverdyck, & Tamir, 2019b). This suggests that part of the reason that neural representations of current states resemble neural representations of likely future states may be because all states are encoded as coordinates within the mental state map.

Conclusion

The brain makes sense of other people's minds by charting conceptual maps of social stimuli, such as actions, mental states, and traits. These maps make the deluge of information from social world more tractable by reducing the complexity of these stimuli down to coordinates on a few essential psychological dimensions, such as valence or animacy. Moreover, these maps allow people to make accurate, automatic predictions about others' actions and mental states. Just knowing where someone is on these maps can tell you a great deal about where they are going next because shorter journeys are more likely than lengthy ones. The short hop from joy to gratitude is far more likely than the arduous trek from delight to despair. The research outlined in this chapter offers insight into how people make sense of other people in order to navigate the choppy waters of everyday social life.

So far, research has focused on how the most basic information conveyed by social maps—the locations of traits, states, and actions—can inform social prediction. However, as with physical maps, social maps can also convey other forms of information. For example, if you look at oceans on many physical maps, you will see small arrows that indicate the direction of prevailing winds and currents. If you dropped a sealed bottle in the ocean at given location, you could use its location, along with knowledge about local water currents to make a precise—and directional—prediction about its future location. The space of mental states likewise has prevailing winds and currents: there are statistical regularities in the trajectories that states and actions follow on their respective maps. For example, people in high energy states are likely to gradually flow toward lower energy states as they tire themselves out. Future research must attempt to map these vector fields in the social domain to further refine the predictive framework we describe here.

The framework for predictive social cognition described in this chapter faces at least four additional challenges (Saxe, 2018). First, the types of social knowledge mapped so far all relate to the person—what they are doing, how they are feeling, and their character. However, one of the most potent drivers of real-world behavior does not dwell within any one person. Instead, the situation, or context, crucially shapes how one will think, feel, and act. The current framework must expand to incorporate the power of the situation and its role in social predictions. Fortunately, in recent years, behavioral research and text analysis have suggested potential maps of the situation layer (Parrigon, Woo, Tay, & Wang, 2017; Rauthmann et al., 2014). Future research may productively test how well such maps explain neural representations of situations, and whether these maps have the same predictive properties as the maps of action and state layers described above.

Second, cultural differences may shape the way people construct maps of social knowledge. Societies differ greatly in the way they perceive emotions, the value they place on different traits, and in the actions that people typically perform (Ching et al., 2014; Gendron, Roberson, van der Vyver, & Barrett, 2014; Tsai, Knutson, & Fung, 2006; Watson-Jones & Legare, 2016). Measuring the generalizability and variability across cultures of the dimensions identified above must be another priority for this research program.

Third, although the model of predictive social cognition described in this chapter has demonstrated its ability to predict real social experience, it faces another major challenge in making these predictions precise. Specifically, this probabilistic framework must incorporate propositional information. For instance, the current framework can describe the properties of "desire" as an abstract mental state, but the meaning of desire can change depending on what one desires. Desiring a cheeseburger and desiring a job are both recognizable forms of desire, but each would predict dramatically different behaviors. Other models of theory of mind—such as Bayesian inverse planning models (Baker, Saxe, & Tenenbaum, 2009)—can deal well with these sorts of propositional problems. However, these types of models do not scale to real-world experience as easily as the current model. Finding ways to unite these models may prove challenging, but a variety of emerging methods, such as word vector embeddings to quantify semantics (Pennington, Socher, & Manning, 2014), may help address these challenges, as similar neural networks have shown to also implicitly represent propositional relations (McCoy, Linzen, Dunbar, & Smolensky, 2018).

The fourth major challenge faced by the predictive model of social cognition is the question of how the mind learns to map the social world. The dimensions of mental state representation arise over the course of development—infants do not start off understanding mental states on all of the dimensions which adults do (Nook, Sasse, Lambert, McLaughlin, & Somerville, 2018). Do children learn new dimensions by observing statistical regularities in emotion dynamics? Perhaps—it is well known that children possess the ability to learn the transitional probabilities between components of speech (Saffran, Aslin, & Newport, 1996), so the same might well be true with respect to other social stimuli. However, it is also possible that children may have "built-in" core knowledge, or inductive biases that help them learn social maps more adeptly than they otherwise might. Future developmental and comparative research, as well as study of machine intelligence, may help to answer this question.

As the model of predictive social cognition described in this chapter becomes more comprehensive and more refined, it holds considerable promise for addressing problems of societal importance. For instance, it may provide a concrete way to quantify abnormal social cognition, such as in Autism Spectrum Disorder, or to track how children learn social concepts over development. Indeed, a theory of "Mind-space"—similar to the mental state and trait layers of our predictive model—has been proposed as a way to gain traction on individual differences in social cognition (Conway, Catmur, & Bird, 2019). Predictive social cognition may also provide a roadmap for enhancing artificial intelligence in the social domain, allowing smart devices to better anticipate people's needs and to interact with people in more natural, human-like ways. Finally, the shortcuts the brain takes to social understanding could reveal the precise sources of harmful social biases and suggest potential approaches to mitigating them. Even optimistically, such development remains many years away, but nonetheless, understanding predictive social cognition holds much promise for the future.

References

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition, 113*(3), 329–349.

Ching, C. M., Church, A. T., Katigbak, M. S., Reyes, J. A. S., Tanaka-Matsumi, J., Takaoka, S., … Rincon, B. C. (2014). The manifestation of traits in everyday behavior and affect: A five-culture study. *Journal of Research in Personality, 48*, 1–16.

Conway, J. R., Catmur, C., & Bird, G. (2019). Understanding individual differences in theory of mind via representation of minds, not mental states. *Psychonomic Bulletin & Review*, 1–15.

Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology, 40*, 61–149.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*(3), 223–241.

Fiske, S., Cuddy, A., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878–902.

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion, 14*(2), 251–262.

Goldberg, L. R. (1990). An alternative "description of personality": The big-five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216–1229.

Hohwy, J., Roepstorff, A., & Friston, K. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition, 108*(3), 687–701.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review, 93*, 1449–1475.

McCoy, R. T., Linzen, T., Dunbar, E., & Smolensky, P. (2018). RNNs implicitly implement tensor product representations. *ArXiv, 1812*, 08718.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81–90.

Mitchell, J. P. (2008). Contributions of functional neuroimaging to the study of social cognition. *Current Directions in Psychological Science, 17*(2), 142–146.

Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2018). The nonlinear development of emotion differentiation: Granular emotional experience is low in adolescence. *Psychological Science, 29*(8), 1346–1357.

Parrigon, S., Woo, S. E., Tay, L., & Wang, T. (2017). CAPTION-ing the situation: A lexically-derived taxonomy of psychological situation characteristics. *Journal of Personality and Social Psychology, 112*(4), 642.

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543).

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*(1), 79–87.

Rauthmann, J. F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C. S., Sherman, R. A., … Funder, D. C. (2014). The situational eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology, 107*(4), 677–718.

Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science, 17*(1), 31–35.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928.

Saxe, R. (2018). Seeing other minds in 3D. *Trends in Cognitive Sciences, 22*(3), 193–195.

Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology, 84*(5), 1041–1053.

Tamir, D. I., & Thornton, M. A. (2018). Modeling the predictive social mind. *Trends in Cognitive Sciences, 22*(3), 201–212.

Tamir, D. I., Thornton, M. A., Contreras, J. M., & Mitchell, J. P. (2016). Neural evidence that three dimensions organize mental state representation: Rationality, social impact, and valence. *Proceedings of the National Academy of Sciences, 113*(1), 194–199.

Tarhan, L., & Konkle, T. (2019). Sociality and interaction envelope organize visual action representations. *BioRxiv*, 618272.
Thornton, M. A., & Mitchell, J. P. (2017). Consistent neural activity patterns represent personally familiar people. *Journal of Cognitive Neuroscience, 29*(9), 1583–1594.
Thornton, M. A., & Mitchell, J. P. (2018). Theories of person perception predict patterns of neural activity during mentalizing. *Cerebral Cortex, 28*(10), 3505–3520.
Thornton, M. A., & Tamir, D. I. (2017). Mental models accurately predict emotion transitions. *Proceedings of the National Academy of Sciences, 114*, 5982.
Thornton, M. A., & Tamir, D. I. (2019a). People accurately predict the transitional probabilities between actions. *PsyArXiv.*
Thornton, M. A., & Tamir, D. I. (2019b). Six dimensions describe action understanding: The ACT-FASTaxonomy. *PsyArXiv*.
Thornton, M. A., & Tamir, D. I. (2020a). People represent mental states in terms of rationality, social impact, and valence: Validating the 3d mind model. *Cortex, 125,* 44–59.
Thornton, M. A., & Tamir, D. I. (2020b). Perceiving actions before they happen: Psychological dimensions scaffold neural action prediction. *Social Cognitive and Affective Neuroscience.*
Thornton, M. A., Weaverdyck, M. E., Mildner, J. N., & Tamir, D. I. (2019). People represent their own mental states more distinctly than those of others. *Nature Communications, 10*(2117).
Thornton, M. A., Weaverdyck, M. E., & Tamir, D. I. (2019a). The brain represents people as the mental states they habitually experience. *Nature Communications, 10*(2291).
Thornton, M. A., Weaverdyck, M. E., & Tamir, D. I. (2019b). The social brain automatically predicts others' future mental states. *Journal of Neuroscience, 39*(1), 140–148.
Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology, 90*(2), 288–307.
United States Bureau of Labor Statistics. (2003). *American time use survey*. Washington, DC: Author.
Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage, 48*(3), 564–584.
Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music–brain responses to rhythmic incongruity. *Cortex, 45*(1), 80–92.
Watson-Jones, R. E., & Legare, C. H. (2016). The social functions of group rituals. *Current Directions in Psychological Science, 25*(1), 42–46.
Weaverdyck, M. E., Thornton, M. A., & Tamir, D. I. (2020). The neural geometry of mental state representation is stable across modalities and targets. Manuscript in preparation
Zhao, Z., Thornton, M. A., & Tamir, D. (2018). Accurate prediction of emotion transitions is associated with social benefits. *PsyArXiv*.