Pluralism in Social Cognition and Predictive Processing

Elmarie Venter (Elmarie.Venter@rub.de)
Institut für Philosophie II, Universitätsstr. 150, 44801
Bochum. Germany

Abstract

In this paper, I explore two issues with the pluralist approach to social cognition. First, the pluralist approach does not assume any particular cognitive framework that could accommodate the variety of strategies in social cognition. Second, the pluralist approach suggests that a variety of strategies are employed in social cognition but neglects to address how mediation takes place between strategies. I argue that both these issues can be addressed if the pluralist approach situates itself in the predictive processing framework. To elaborate on this, I propose that 1) the strategies for social cognition include obtaining and testing theories in generative models about the behavior and mental states of others, 2) interactional synchrony is a strategy employed in simple social situations and 3) affordances play an unprecedented role in mediating between strategies.

Keywords: social cognition; pluralism; predictive processing

Introduction

Interacting with other agents is a central part of our everyday lives. We rely on others for many aspects insofar as other agents offer useful information about the world and influence our possibilities for action and interaction. For this reason, it is important that we, as agents, have the capacity to predict and explain the behavior of others. Despite the importance of the social domain and many theoretical proposals about how understanding each other works, there is relatively little agreement about the mechanisms and processes that enable and constitute social cognition (Baron-Cohen et al, 2013; Carruthers & Smith, 1996). The past few decades have seen the development of theory-theory, simulation-theory, two-systems theory and interaction theory to name just a few. Ultimately, these approaches have one thing in common: a single default strategy is used for how we predict and explain the behavior of others.

Opting for a single strategy means counterexamples always pop up that the relevant strategy cannot explain. Therefore, we may need to look to pluralist approaches which adopt a 'no default procedure' approach and instead posit a variety of strategies that are employed in social situations. Though promising, the pluralist approach still lacks a cognitive framework in which the strategies are realized. To address this concern, I first delve into the social cognitivist's toolbox and explore the various strategies posited in social cognition to motivate the move towards a pluralist account. I then further develop the pluralist approach in the predictive processing framework and propose that social cognition consists of a variety of strategies that are employed depending on the situational context. Particularly that the strategies available to agents include generating models about the behavior and mental states of other agents, and interaction regulated through interactional synchrony. I then tackle a second issue with pluralism, namely addressing how mediation between strategies take place by proposing that mediation between strategies is driven by the affordances available to the agent in the social situation.

The social cognition toolbox

For the sake of brevity, I only briefly sketch out the commitments of the major players in the social cognition debate and highlight a few disagreements.

Mindreading Theories

Over the last several decades, the debate has been largely shaped by theory-theory (TT). One of the primary claims of TT is that children form theories about behavior and revise these theories in light of new data. This equips children with a theoretical toolbox to explain and predict the behavior and mental states of other agents. It provides a promising developmental account insofar as radical theory-shifts can be observed in children holding theories that are based on perception and desire (early infancy) to holding theories that are based on an understanding of the notion of a belief (perhaps from the age of four). It thus provides a good explanation of how children continuously learn about behavior of other agents and how to shape their own behavior. Still there are some disagreements in the debate with regards to the developmental aspect with Gopnik and Wellman (1992), on the one hand, suggesting that infants are psychologically aware of perceptions and desires and only around age four or five children develop a fully-fledged representational model of the mind when "almost all psychological functioning is mediated by representations" (Gopnik and Wellman, 1992). On the other hand, scholars like Carruthers (2013) and Scholl and Leslie (1999) propose that explaining and predicting behavior is based on innate, domain-specific principles present from birth.

TT is highly dependent on a representational account of the mind and though this may not be a problem in and of itself, it posits a rather complex and sophisticated process to understand and interact with other agents. Mental-state information is processed slower than perceptual information and may not become incorporated in the perceptual state – this is an issue identified by Carruthers himself (Carruthers, 2015). The issues with TT have led to

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

the development of theories that posit more efficient strategies; strategies that do not require constructing elaborate theories about the mental states of other agents.

One of those alternatives is simulation theory (ST): the view that we ascribe mental states to other agents by imagining ourselves in their situation. Through reenacting and mimicking the situation and generating possible future states (based on past experiences), we can predict the behavior of other agents. Consider a popular example discussed by Goldman (1989). In the example, while playing a game of chess, subjects use their own mental states to model their opponent's mind and thereby determine what has taken or will take place.

The example illustrates the core ideas of ST, but also its weaknesses. In my attempt to understand the decisions that my opponent in a game of chess will make, I simulate her mental states in my own mind. More often than not, however, these simulations seem to be inaccurate due to limitations yielded by cognitive biases, social norms, past experiences, and different cultural backgrounds. Perhaps an elaborate theory that mediates these factors could help out in this situation? For this reason, many scholars have opted for a hybrid of TT and ST (see e.g., Nichols and Stich, 2003).

However, even when combined into a hybrid account, TT and ST face challenges. Both accounts are committed to the idea that the mental states of other agents are hidden and must therefore be mediated and interpreted either through inference or simulation – both strategies requiring elaborate representation of my own and the other agent's mental landscape. As previously mentioned, appealing to mental representations is not necessarily a weakness but these accounts and their cousins (i.e., two-systems theory and modularity theory) posit an epistemic gap between mind and behavior, and that creates a plethora of issues that must be addressed. This has led to the development of an alternative view that has gained popularity also in light of the development of embodied and enactive cognition.

Direct Perception Theories

The relation between direct social perception and embodiment comes in various strengths. First, it can be said that *the mind is perceptually co-present within bodily expressions* (Krueger, 2018). To illustrate this, consider the following situation. When perceiving a car, I see only the part of the car that is facing me – say the driver's side. The passenger's side of the car is hidden from me but I nevertheless experience the car as a three-dimensional object. The hidden sides of the car are perceptually co-present in a phenomenological sense. This can be applied to how we perceive other agents. When we see the bodily expression and behavior of another agent, we also experience the associated mental state. Smith (2010) writes:

> "Just as the rear aspect of the book is visually present without being visually presented, so another's misery is visually present even though only their frown is visually presented."
>
> (Smith, 2010)

Although this account takes a step away from the mediation accounts in the previous section, it is still compatible with the idea that mental states are hidden in some sense. We experience the mental states of others only through seeing their behavior and only have indirect access to their beliefs, desires and intentions.

The second way in which the direct social perception theory can be construed is by characterizing the relation between mental states and behavior in terms of *constitution* rather than mere co-presentation. On this approach, the gap between mind and behavior is entirely dissolved and "[s]ome mental states are concretely embodied within the expressive behavior we see" (Krueger, 2018). When I perceive another agent, I literally *see* her mental states, not just the causal effect of her mental states. The constitutive approach to direct social perception recognizes that minds consist of internal and external phenomena — denying internal realization would be implausible. Nevertheless when perceiving other agents, we have direct perceptual access to the parts of the mind that are externally realized (Krueger, 2018). A consequence of this approach is that rather than positing a first- or third-person stance in a social interaction, two first-person perspectives become integrated into a second-person perspective. The second-person stance is grounded in the following idea: when two social partners *interact*, they know one another in a way that is different from merely observing the other (Reddy, 2008). To explore this idea further, let us take a brief look at interaction theory (IT).

Gallagher (2001) offers interaction theory (IT) as an alternative account to theory-theory and simulation theory. A key feature of IT is that it defines mental states as expressed in our embodied actions and visible to other agents. In this sense, mental states are not hidden phenomena that require interpretation and there is no epistemic gap between mind and behavior. Instead, our understanding of other minds is primarily enabled through *embodied practices* that are emotional, sensorimotor, perceptual and non-conceptual, and thereby not based on inference or internal simulation (Gallagher, 2001). For example, my desire to be close to my partner is expressed in my embodied action as I stand closer to him and hold his hand. This may also express my belief that showing affection is a way to get closer to someone. When we interact with another agent, evaluative understanding – our capacity to know what someone means or intends and how we should respond in particular contexts (Gallagher, 2001) – is the primary strategy employed and only when this breaks down, do we appeal to theories about others or do we simulate their mental states.

Despite the disagreements amongst all the theories discussed above, they have one thing in common: they posit a single dominant strategy for predicting and explaining the behavior of others. Each of the approaches has strengths but the downfall of each, respectively, is that some counter argument can be presented because the approach is too committed to a single strategy that explains social cognition. There seems to be at least two options around this problem. Either we can attempt to integrate the

various theories and develop something resembling a hybrid theory. This has in fact been done, at least in terms of simulation theory and theory-theory, and arguably two-systems theory is a hybrid between constructivism and nativism. The worry with hybrid integration is that more often than not, theories have commitments that are simply not compatible with others.

Pluralism

Countering the single-strategy and hybrid approaches is the pluralist approach which opts for a variety of strategies employed in social cognition depending on situational context. Fiebich et al. (2016) defines it as follows: "a genuinely pluralist approach holds that social cognition involves a combination of capabilities or processes, any "one of which may be appropriate or practical for one kind of situation, and another for a different kind of situation," but there is no default procedure that we tend to use in every situation (Fiebich & Coltheart, 2015; Gallagher, 2015)". Pluralism in social cognition is not committed to a single strategy for predicting and explaining the behavior of others. Nor does it posit that any strategy is primary or more dominant than any other. Given the diverse nature and complexity of social cognition, it thus seems like the most promising way forward. Also especially in light of the concerns raised against all the other theories to date.

Despite the strengths of the pluralist approach, it is still in its infancy and requires further development. I identify two issues to tackle for an adequately developed pluralist account of social cognition. First, the pluralist approach does not posit any particular cognitive framework that could accommodate the variety of strategies employed in social cognition. Mindreading theories – according to which mental states are hidden and thus only accessible to other agents through inference – assume cognitivism/computationalism as their foundation. This leads to a commitment to mental representations but at the same time a neglect of the embodied and enactive aspects of social cognition. On the other hand, direct perception theories are more committed to enactive approaches and thereby grant less importance to mental representations, i.e., it assumes that mental states are directly perceivable and thereby not hidden from others. Though both of these underlying frameworks have validity as cognitive theories, pluralism does not seem committed to either. A framework is thus desired but importantly one that can accommodate both these underlying assumptions. I do not assume that these commitments are mutually exclusive; it is possible to accept that sometimes mental states are hidden but other times we can directly perceive the states of another agent. Predictive processing, as I will defend it here, accommodates this and thus offers a home for pluralism to be hosted in. Second, the pluralist approach suggests that a variety of strategies are employed in social cognition but neglects to say how mediation between strategies take place. What determines which strategy is employed? What conditions need to be in place for an agent to make use of a theory rather than relying on the interaction to regulate the social situation? These are all unanswered questions. I will propose that, through presenting certain affordances to the agent, the context mediates the employment of strategies.

In the next section, I explore how situating pluralism in predictive processing can offer a promising solution to these issues.

Predictive processing and social cognition

One of the strengths of the predictive processing account (PP) is that, though it sets out to explain perception and action, it can be extended to a variety of cognitive phenomena including social cognition. Broadly construed, there are two radical extremes of PP that dominate the literature: cognitivist PP (Hohwy, 2013) and free energy enactivism (Bruineberg et al, 2018). It is, however, possible to get the best of both worlds when one adopts a more embodied approach to PP (see e.g., Venter, 2021; Clark, 2015, 2017). Such an approach maintains the representational nature of cognition while at the same time emphasizing the embodied, enactive nature of our interactions with the world (Venter, 2021). As I will show in the following subsections, such an account of PP ideally complements the pluralist approach in social cognition and avoids many of the shortcomings of the radical interpretations of PP as well as the single strategy accounts of social cognition. The radical interpretations of PP fall victim to the same problems faced by traditional accounts of social cognition insofar as they propose a primary strategy for coping with social situations. I do not set out to develop an elaborate account of PP in this paper; this has been done elsewhere (see Venter 2021; 2024). Instead, I briefly present the main tenets of PP that will be relevant for further discussion of pluralism in social cognition.

The predictive processing toolbox

PP turns the standard computationalist story on its head and is rooted in the idea that perception is driven by top-down processes. As such, the system is "constantly trying to guess the present" (Clark, 2017). Top-down predictions are generated and then matched with incoming (bottom-up) information. The function of top-down predictions is thus to "explain away" incoming information leaving only prediction errors to be propagated within the system (Clark, 2013). This has the upshot of not encountering a bottleneck of information processing because only prediction errors are propagated in the system. This process occurs at many different levels in the processing hierarchy where, at each level, the system is trying to predict its own states. Higher level predictions involve more abstract and temporally extended states, and lower level predictions process more fine-grained states. Simply put, the brain is in the business of finding the best set of predictions based on precision and accuracy to explain the sensory input given the context and prior knowledge. This is encoded in generative models (Clark, 2013). In other words, generative models encode activity from lower levels in the hierarchy in ways that enable better predictions about the world given prior knowledge and context (Clark, 2013).

But this is only half the story. Prediction error minimization (PEM) is not the only objective of the system. Though constructing and updating theories about the world (and other agents) is a necessary part of cognitive processes, it certainly is not the only function. The individual acts upon the world to regulate vital bodily parameters but also to adapt and react to changing environmental contexts. This is not always done through updating generative models which is a cognitively costly transaction in light of an ever-changing dynamic environment. Inferring the hidden causes of input is less important in certain situations. Instead preparing the right actions and reactions in response to the changing environment is more efficient and reflect the nature of the system more adequately (see e.g., Quiqley et al, 2021). This captures the action-oriented nature of predictive processing (Clark, 2017). For social cognition, this means that an agent acts in a social situation by directly perceiving the cues (i.e., facial expressions, gestures, etc.) and – depending on the environmental context and past conditioning – acts without mediation by hypotheses about the situation. Timely adjustments in behaviour thus maintain the interaction which is regulated by the situation and the bodily states of the agents. Social cognition emerges as a situated, context-dependent response that reflect the whole embodied agent's best estimate of anticipated future states. Thus, some situations do not rely on inferences as discrete mental entities but rather on dynamic, continuous processes shaped by individual differences, cultural influences, and situational factors.

Agent-models

A key feature of predictive processing is the deployment of generative models in our perception of and engagement with the world. Given that other agents are an integral part of the world in which we are embedded, we can expect that the predictive system also generates models of other agents. Hohwy & Michael (2017) call these "agent-models" and propose that we construct and update these models to predict the behavior of other agents and to shape our own behavior. In a recent article, de Bruin & Michael (2018) propose that on the predictive processing approach, other agents are considered to be a cause of sensory input on a par with other objects in the world. They write:

> "Just as it is sometimes helpful in reducing uncertainty about the behavior of edges and surfaces and colors to postulate teacups as the bearers of those lower-level features, it is also sometimes useful to postulate other people."
>
> (de Bruin & Michael, 2018)

This does not suggest that other agents are perceived as mere objects. Instead, what de Bruin and Michael are hinting at is that the predictive system has generative models of agents that generate predictions about the movements and intentions of other agents at various levels in the hierarchy in the same way as it does for ordinary objects.

Agent-models are active even at very low levels of the predictive hierarchy where the detection of faces, gaze direction and emotions trigger higher level predictions that are relevant to agents with intentions. To illustrate this, consider how the predictive processing account explains the phenomenon of pareidolia (i.e. to see faces in ordinary objects). When detecting human-like features such as something resembling eyes and a mouth, a whole cascade of predictions is triggered. Generative models about the structure of faces activate models that predict the emotions expressed by the face and other qualities of human agents. For example, human bodies are subject to particular biomechanical laws. Generative models about these biomechanical laws are activated when we perceive another agent and predictions about the movement and intentions of agents are generated. There is evidence to support that when we perceive movements that violate biomechanical laws, there is greater response in the brain areas that are associated with biological motion (Costantini et al., 2005). This suggests that prediction error is processed for unusual biomechanical movements.

Agent-models are also used in more complex social situations. Human agents are understood to be agents that behave according to situational regularities and their behavior is governed by various norms and conventions (de Bruin & Michael, 2018). It would, for example, be very unlikely that someone ignores a queue in the supermarket and simply walks past numerous people to place their items on the conveyer belt for payment. We are taught to be sensitive to these norms from a very young age and come to behave according to norms through imitating and learning from other agents. A child visiting the supermarket with her dad and seeing someone reproached for skipping the line would learn that this behavior is not acceptable. We also tend to enforce norms and conventions by correcting behavior that is not in line with situational contexts. This can be understood as minimizing prediction error through active inference (de Bruin & Michael, 2018).

Generative models about agents also generate predictions about psychological features of particular agents. These predictions involve regularities that persist over longer time scales, such as preferences, character traits, and beliefs, and enable us to make predictions about the behavior of agents. For example, if my friend has a preference for hot chocolate (because she always orders this when we meet), I would be surprised when we meet and she frowns in disgust at my hot chocolate. If I have a model of the beliefs and desires of another agent, I can make predictions about their behavior. For a lower-level processing example, consider the false belief task (see Wimmer and Perner, 1983). In false belief task experiments, children tend to look longer at the scene when the expectation of where the agent will look for a moved item is violated (Baillargeon et al., 2010). This suggests that children hold predictions about the behavior of another agent which, when violated, results in longer looking times to gain more information that explains the outcome. The notion of active inference can provide some insight here too. In active-helping experiments, children directed agents to the location of the desired toy and extracted it for the agent when their expectations were violated (Buttelmann et

al., 2009) which suggests that children actively reduce prediction error when their expectations are violated.

Positing generative models is a necessary component of the predictive processing account but this cannot be all there is to an explanation of how we perceive and understand other agents. Other strategies may be more efficient or effective depending on the situational context and its complexity or simplicity, and importantly it is sometimes cognitively costly to rely on generative models. In some situations, embodied interaction is the more efficient strategy for interaction. Such embodied interactions could, of course, be represented or encoded in the models but this point has received much criticism because it does not seem to adequately capture our experience of social interactions. Consider the three features that characterize our social interactions. First, we engage with the movements and gestures of our social partners. For example, if someone points at something, we look at it and if someone extends their hand, we return a handshake. In this sense the body plays a constitutive role in social cognition in virtue of embodied interactions rather than the body being merely available in terms of interpretation and inference. Such embodied interactions manifest in an action-reaction loop which will be unpacked in terms of intercorporeality and interactional synchrony below. Second, as suggested by direct social perception, we directly perceive the mental states of other agents in their bodies. The other agent need not be differentiated into internal and external dimensions, and their mental states are not always hidden from us. Third, our social interactions always appear in context. This feature is considered by the agent-model theory of social cognition but ordinary situational interactions do not necessarily require appeal to high level abstract generative models. Though, we may need to appeal to theories about culture and norms when we enter a novel situation. This will be unpacked in the section on affordances as a way of mediating between what strategy is adequate in a given setting.

A model-only based account thus cannot fully explain social cognition for two reasons. First, it construes the mental states of others as always hidden and not observable which requires the use of theories and models even at very low levels in the hierarchy. In doing so, the account offers only generative models as the sub-personal mechanisms that realize social cognition and thus also neglects to consider the distinct and unique mechanisms involved in self- and other-processes. One such mechanism is the body and how it is experienced 'from the inside'. By granting the body a constitutive role in realizing social cognition, a mechanism is identified that regards self-processes as different to how the mental states of other agents are processed. When there is alignment between action and intention, the movements are processed as belonging to the subject identified in virtue of the minimal self and bodily intentionality. Second, it claims that the primary strategy in our attempts to understand and predict the behavior of other agents is the use of theories and models. This claim neglects to consider that most of our social interactions do not require interpretation of mental states (of self or other) but rather our interactions are direct and embodied. To address these two concerns and support the claim that the body has a constitutive role in social cognition, I appeal to the phenomenological notion "intercorporeality" and explain this at the sub-personal level of mechanisms in terms of interactional synchrony— a term borrowed from psychology (Bernieri & Rosenthal, 1991) and also applied by free energy enactivism (Friston & Frith, 2015).

Intercorporeality and interactional synchrony

Intercorporeality, like direct social perception, abandons the mind-behavior dualist assumption according to which mental states are only accessible to the subject. Instead, mental states are understood to be expressed (and thus observable) in our bodily expressions. In terms of interaction, we engage in a reciprocal relationship between our and other bodies and can thereby directly grasp the intention of another's action. This approach is grounded in the idea that we interact with others through the relation between the bodies of self and other.

We perceive the actions and intentions of the other through our body in a pre-reflective manner. Merleau-Ponty relies on the example of an infant opens his mouth when his caregiver playfully takes one of his fingers to pretend to bite it. "Biting" has an intersubjective signification (Merleau-Ponty, 2012).

In this scenario, the baby pre-reflectively acknowledges the caregiver's intention through his body. This strategy suggests that we understand others by grasping their actions through our own body and reciprocally finding our own possibility of action in another's body (Tanaka, 2015). Despite the similarities, intercorporeality can be differentiated from simulation theory in that instead of passively receiving sensory input, simulating the state, and acting upon it, we are set out to actively seek potential actions. Intercorporeality is thus not only characterized by mirroring the behavior of others, it also involves interaction through action and reaction which follows the perception of affordances.

There is, however, still one point of contention in that intercorporeality, as a phenomenological concept, only provides explanation at the personal level and is not concerned with the sub-personal mechanisms that may realize the phenomenon. To fill this gap in explanation, interactional synchrony as applied in free energy enactivism can provide the sub-personal mechanisms of explanation.

Interactional synchrony occurs during social interactions when two agents synchronize their rhythms and movements (Bernieri et al., 1988). Infants, for example, synchronize bodily movements to the speech patterns of their caregivers and adults tend to match the flow of gestures and movements to the speech pattern of their interlocutor (Feldman, 2007). In the context of the predictive processing account, synchrony can be understood as social interaction partners staying within a particular range of states determined by the context. Such interpersonal coordination occurs when one person produces a meaningful reaction based on their social

partner's action and behavior. Suppose we are sitting in a busy coffee shop having a conversation, you wish to tell me something about someone standing behind me but rather than project the statement loudly, you whisper it. The natural response, on my part, is to lean closer to you to hear what you are saying and in realizing of whom you are speaking, I shift my body in a way that brings this person into my view. The reaction to you whispering causes a subsequent action and the feedback loop continues. Perhaps as I turn around, you naturally pull my arm back so as to prevent me from turning around completely and drawing attention to our conversation.

In a paper on synchronization, (Friston & Frith, 2015) use the formal notion of *generalized synchronicity* to explain communication between two predictive systems. I suggest that these two notions, interactional synchronicity and generalized synchronicity, capture the same interpersonal coordination involved in interaction though interactional synchrony requires more than mere synchronization. In the study Friston & Frith (2015) conducted, simulations of birdsong were used to illustrate the omission-related responses involved in communication. The synthetic 'birds' were in a scenario where they could hear themselves and the other sing. Each bird listened for two seconds and sang for two seconds following the coordinated behavior that when one bird sang, the other listened (Friston & Frith, 2015). When placed out of earshot from one another, their expectations followed independent trajectories. But when they moved closer together so that they could hear one another, they synchronized almost immediately. The results show that one bird is not simply mirroring the other but rather that one reciprocally acts following the action of the other.

The basic idea illustrated by the above study is that we can interact with one another because we have a capacity for action and attenuation in interpersonal communication. In any given context, there is a relative set of states in which social partners can successfully interact. This set of states enforces generalized synchrony in that the states of the other imposes constraints on the state that the individual can occupy (Friston & Frith, 2015). This is based on the embodied perception of the other's body in action as well as the situational context. As previously mentioned, the perceptual process is not one of passively receiving information, processing it, and then acting on the environment. Instead, it is a process of using our embodied skills to explore possibilities of action in the environment. Interactional synchrony thus arises, not out of mediated interpretations of each other's behavior but, out of embodied perception of the interaction partners (Tanaka, 2015). Interactional synchrony and generative models are two ways in which we can understand other minds. These two strategies for understanding others gets us minimally closer to a fully developed pluralist account of social cognition. Underlying the details of how these two strategies are mediated in social contexts, is the notion of affordances.

Affordances

Affordances offer the key to mediating between strategies in social contexts. In the context of social interaction, the behavior of other agents is perceived as affording a reaction and, in this sense, other agents present us with social affordances. Social affordances are possibilities for interaction offered by other agents in the environment. It is important to note that such social affordances are not presented in isolation. The environment presents us with a landscape of affordances at many different timescales ranging from a long-term timescale involving cultural and societal norms to a short-term timescale involving contextual interaction and reciprocal behavior.

The environment and other agents can thereby afford direct reciprocation because the social cues are directly perceivable. For example, a person waiting in line with me at the supermarket affords a conversation, a couple in closed conversation does not, an extended hand affords a handshake and someone who looks angry and has their arms crossed does not. On the other hand, the social context could be more complex and thereby require inference of the mental states of another agent in which case appealing to models or theories of behavior, cultural background, etc. would be more efficient and effective. Consider, for example, if I see a couple having an argument across the street. To understand what is happening, I may appeal to previous experiences and theories I have about why and when someone may have an argument in public. The situation could, however, quickly change. If one of the parties get physical, I may directly respond and intervene because the situation affords interaction. I no longer have to infer the argument but can instead directly perceive the situation which provides certain possibilities for action: either the situation elicits intervention or to ignore the situation and go about my day. In summary, social situations can offer a variety of possible responses that require flexibility in terms of strategy employed. The possibility for action ultimately determines which strategy is employed.

Conclusion

In considering the strategies employed in social cognition, the predictive processing account suggests that we employ generative models to predict and explain the behavior of others. At the very least, however, theories are employed in only some instances – perhaps more complex social situations. We can expect that in most instances of social cognition interactional synchronization suffices. That means we often interact with other agents without consciously reflecting on our behavior and without the need to infer the mental states of the other. The context, and the affordances it makes available to the agent, mediates between the strategies are ultimately determining which strategy is to be employed.

References

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, 14(3), 110–118.

Baron-Cohen, S., Tager-Flusberg, H., & Lombardo, M. (2013). *Understanding other minds: perspectives from developmental social neuroscience*. Oxford University Press.

Bernieri, F. J., Reznick, J. S., & Rosenthal, R. (1988). Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother-infant interactions. *Journal of Personality and Social Psychology*, 54(2), 243–253.

Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: behavior matching and interactional synchrony. In R. S. Feldman & B. Rimé (Eds.), *Studies in Emotion and Social Interaction. Fundamentals of Nonverbal Behavior* (pp. 401–432). Cambridge University Press.

Bruineberg, J., Kiverstein, J. and Rietveld, E. (2018). The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. *Synthese*, 195(6), pp.2417-2444.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112(2), 337–342.

Carruthers, P., & Smitth, P. (1996). *Theories of theories of mind*. Cambridge University Press.

Carruthers, P., 2013. Mindreading in infancy. *Mind & Language*, 28(2), pp.141-172.

Carruthers, P. (2015). *The centered mind: What the science of working memory shows us about the nature of human thought.* OUP Oxford.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204.

Clark, A. (2015). Radical predictive processing. *South. J. Philos.* 53, 3–27.

Clark, A. (2017). How to Knit Your Own Markov Blanket: Resisting the Second Law with Metamorphic Minds. *Philosophy and Predictive Coding*, 1–19.

Costantini, M., Galati, G., Ferretti, A., Caulo, M., Tartaro, A., Romani, G. L., & Aglioti, S. M. (2005). Neural systems underlying observation of humanly impossible movements: An fMRI study. *Cerebral Cortex*, 15, 1761–1767.

de Bruin, L., & Michael, J. (2018). Prediction error minimization as a framework for social cognition research. *Erkenntnis*, 1–20.

Feldman, R. (2007). Infant biological foundations synchrony and developmental outcomes. *Current Directions in Psychological Sciences*, 16(6), 340–345.

Fiebich, A., & Coltheart, M. (2015). Various Ways to Understand Other Minds. Towards a Pluralistic Approach to the Explanation of Social Understanding. *Mind and Language*, 30(3).

Fiebich, A., Gallagher, S., & Hutto, D. D. (2016). Pluralism, interaction, and the ontogeny of social cognition. In J. Kiverstein (Ed.), *Routledge Handbook of The Philosophy of the Social Mind*.

Friston, K., & Frith, C. (2015). A duet for one. *Consciousness and Cognition*, 36, 390–405.

Gallagher, S. (2001). The practice of mind: Theory, simulation or primary interaction? *Journal of Consciousness Studies*, 8(6–7), 83–108.

Gallagher, S. (2015). The new hybrids theories of social cognition. *Consciousness and Cognition*.

Goldman, A. (1989). Interpretation Psychologized. *Mind and Language*, 4(3).

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language*, 7(1–2), 145–171.

Hohwy, J. (2013). *The predictive mind*. OUP Oxford.

Hohwy, J., & Michael, J. (2017). Why should any body have a self? In F. de Vignemont, A. J. Alsmith, R. L. Kosut, & S. Mussa-Ivaldi (Eds.), *The Subject's Matter: Self-Consciousness and the Body*. (pp. 363–391). MIT Press.

Krueger, J. W. (2018). Direct social perception. In A. Newen, L. de Bruin, & S. Gallagher (Eds.), *The Oxford Handbook of 4E Cognition* (pp. 301–320). Oxford University Press.

Merleau-Ponty, M. (2012). *Phenomenology of Perception* (D. Landes, Ed.). Routledge.

Nichols, S. and Stich, S.P. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds.* Oxford University Press.

Reddy, V. (2008). *How Infants Know Minds*. Harvard University Press.

Scholl, B.J. and Leslie, A.M. (1999). Modularity, development and 'theory of mind'. *Mind & Language*, 14(1), pp.131-153.

Smith, J. (2010). Seeing other people. *Philosophy and Phenomenological Research*, 81(3), 731–748.

Quigley, K.S., Kanoski, S., Grill, W.M., Barrett, L.F. and Tsakiris, M. (2021). Functions of interoception: From energy regulation to experience of the self. *Trends in neurosciences*, 44(1), pp.29-38.

Tanaka, S. (2015). Intercorporeality as a theory of social cognition. *Theory & Psychology*, 25(4), 455–472.

Venter, E. (2021). Toward an embodied, embedded predictive processing account. *Frontiers in Psychology*, 12, p.543076.

Venter, E. (2024). Integrating Embodied Cognition and Predictive Processing. In *The Routledge Handbook of Embodied Cognition*, L. Shapiro and S. Spaulding (Eds.). Routledge.

Wimmer, H., & Perner, J. (1983). Belief about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103–128.