

Introduction to debates on embodied social cognition

Shannon Spaulding

Published online: 5 September 2012
© Springer Science+Business Media B.V. 2012

Social cognition, very basically, is our ability to understand and interact with others. In all sorts of everyday situations, we observe, understand, and interact in complex social situations. Following politics, gossiping, playing sports, engaging in pretend play, and driving on a busy freeway are just a few instances of social cognition. The study of social cognition seeks to explain the cognitive architecture of our minds and psychological processes that make it possible for us to engage in social cognition. For much of the past two decades, there has been a two-party debate about social cognition between the *Theory Theory* (TT) and the *Simulation Theory* (ST).

Theory theorists argue that we explain and predict behavior by employing folk psychological theories about how mental states inform behavior. With our folk psychological theories, we infer from a target's behavior what his or her mental states probably are. And from these inferences, plus the psychological principles in the theory connecting mental states to behavior, we predict the target's behavior (Carruthers and Smith 1996; Davies and Stone 1995a; Gopnik and Wellman 1992; Nichols and Stich 2003). In its most general form, the TT holds that mindreading is an information-rich process. The process of understanding others' mental states relies on a rich body of folk psychological information.

Simulation theorists, in contrast, argue that we explain and predict a target's behavior by using our own minds as a model, putting ourselves in another's shoes, so to speak, and imagining what our mental states would be and how we would behave if we were in the target's situation. More specifically, we retrodictively simulate to figure out what the target's mental states could have been to cause the observed behavior, then we take the target's mental states in the form of pretend beliefs and pretend desires as input, run them through our own decision-making mechanism, and take the resulting conclusion and attribute it to the target (Currie and Ravenscroft 2002; Davies and Stone 1995b; Goldman 2006; Gordon 1986; Nichols and Stich 2003). In contrast with the TT, the ST posits an information-poor process.

S. Spaulding (✉)

Philosophy Department, Oklahoma State University, 256 Murray Hall, Stillwater, OK 74078, USA
e-mail: shannon.spaulding@okstate.edu

Understanding others' mental states does not require a large, rich body of folk psychological information. It simply requires an ability to figure out what one would think and do in a target's situation and attribute that to the target.

Though Theory theorists and Simulation theorists disagree over the process underlying the attribution of mental states, they *agree* that how we understand and interact with others in social environments is by explaining and predicting their behavior on the basis of mental state attributions. As such, they agree that the explanation for how we understand and interact with others is what has come to be called *mindreading*. The disagreement is about the cognitive mechanisms of mindreading.

The debate between TT and ST has stagnated recently. Progress has been limited to articulating various hybrid TT–ST accounts. However, within the last 10 years, a new party to the debate has emerged that challenges the mindreading account in both its TT and ST form (Bermúdez 2003; Gallagher 2005, 2009; Hutto 2004, 2008; Lakoff and Johnson 1999). The new account stems from embodied cognition, which is a relatively new research program in cognitive science that challenges cognitivism (Shapiro 2010). Cognitivism has been the dominant view in psychology and philosophy of mind since the 1950s. In fact, it has been so dominant that some have called it the only game in town. Briefly, cognitivism holds that our cognitive capacities should be understood in terms of computational procedures operating on symbolic, internal mental states, and thus, cognitive science should be focused on studying these internal states and processes. Embodied cognition rejects this account of the mind. According to embodied cognition, cognitivism makes the mistake of emphasizing the view of the mind as something to be studied independently of the body and its environment. The emphasis in cognitive science should be on how the body and the environment shape the mind. The embodied cognition argument against mindreading is an application of the more general embodied cognition argument against cognitivism.

Embodied cognition accounts of social cognition, what I shall call Embodied Social Cognition (ESC), aim to explicate how our embodiment shapes our knowledge of others and in what this knowledge of others consists. Although there is much diversity amongst ESC accounts, common to all these accounts is the idea that our normal everyday interactions consist in non-mentalistic embodied engagements. In recent years, several theorists have developed and defended innovative and controversial accounts of ESC. These accounts challenge, and offer deflationary alternatives to, the standard cognitivist accounts of social cognition. As ESC accounts grow in number and prominence, the time has come for a dedicated, sustained debate on ESC and its most controversial and innovative elements. The goal of this special issue is to host such a debate with the aim of bringing clarity to the discussion of social cognition.

The papers in this special issue coalesce into three groups. The papers in the first group critically assess ESC accounts. They argue that ESC is explanatorily inadequate and suggest alternative accounts of social cognition that are more aligned with standard cognitive psychology. Whereas the papers in the first group ultimately argue against ESC and are in favor of positions in the cognitivist camp, the papers in the second group attempt to construct middle-ground positions. These papers identify what they take to be plausible elements of ESC and misguided elements. They argue

that ESC is right to challenge the standard cognitivist picture of social cognition but not right to dismiss every element of it. A common aphorism in these papers is that ESC is guilty of throwing out the baby with the bathwater. The papers in the third group apply ESC to particular domains of social cognition with the goal of showing how useful the account is. They develop an account of ESC and show how it can illuminate some otherwise puzzling and ill-explained aspect of social cognition.

In the rest of this introduction, I shall describe in more detail the papers in this special issue. In doing this, I shall highlight important themes in the papers and discuss the sorts of challenges the arguments in these papers pose for both ESC and cognitivism.

Challenges to ESC

Although several ESC accounts are available, each differing in various ways, all ESC accounts hold that our capacity for social cognition is not based on ascribing mental states to others. Rather, what underlies our ability to understand and interact with others is the capacity for more basic, non-mentalistic, interactive embodied practices. Many of the articles in this special issue assess ESC with respect to its account of developmental psychology. In the developmental psychology literature, the embodied practices that ESC regards as constitutive of social cognition are referred to as “primary intersubjectivity” and “secondary intersubjectivity.”

According to ESC, primary intersubjectivity is the pre-theoretical, non-conceptual, embodied understanding of others that underlies and supports the higher-level cognitive skills posited in the mindreading literature. It is “the innate or early developing capacity to interact with others manifested at the level of perceptual experience—we see or more generally perceive in the other person’s bodily movements, facial gestures, eye direction, and so on, what they intend and what they feel” (Gallagher 2005, p. 204). Primary intersubjectivity is manifested as the capacity for facial imitation and proprioceptive sense of one’s body, the capacity to detect and track eye movement, to detect intentional behavior, and to read emotions from actions and expressive movements of others. Primary intersubjectivity consists in informational sensitivity and appropriate responsiveness to specific features of one’s environment. It does not involve representing those features. It simply requires certain practical abilities (e.g., being sensitive to certain bodily cues), which have been shaped by selective pressures. ESC theorists draw support for primary intersubjectivity from cognitive neuroscience and developmental psychology (Gallagher and Hutto 2007).

The development of secondary intersubjectivity occurs around age 1, and it is marked by a move from one-on-one, immediate intersubjectivity to contexts of shared attention. In addition to tracking eye movement, detecting intentional behavior, and reading emotions, with the development of secondary intersubjectivity, the child develops the capacity to communicate with others about objects and events in the environment. The child’s interactions with caretakers begin to have reference to the things in their environment. At this stage, the child learns to follow gazes, point, and communicate with others about objects of shared attention. With secondary intersubjectivity, the child’s capacity for social understanding is further developed, but according to ESC, this understanding is still non-mentalistic (Gallagher 2005, p. 207).

ESC holds that these embodied intersubjective practices constitute our primary mode of social cognition (Gallagher 2005; Hutto 2008). Daniel Hutto claims, “Our primary worldly engagements are nonrepresentational and do not take the form of intellectual activity” (Hutto 2008, p. 51). Mindreading, it is argued, is a late-developing, rarely used, specialized skill. The embodied practices constituted by primary and secondary intersubjectivity are developmentally fundamental. That is, in order to develop the capacity to have beliefs about others’ mental states, one must first have a grasp of these basic embodied practices. Moreover, embodied intersubjectivity continues to be our principal mode of social interaction even in adulthood.¹ Even as adults, in ordinary circumstances, we do not rely on mindreading to understand others. Our everyday social cognition consists only in these embodied practices and our knowledge of the social norms and behavioral scripts distinctive of our social environments, none of which involves mindreading. To be clear, ESC does not hold that mindreading is impossible. It is simply very rare. We adults use mindreading to understand others only when our primary mode of understanding others—these basic embodied practices—break down.

Thompson

In “Implicit mindreading and embodied cognition,” James Thompson challenges the developmental claims expressed above. Thompson considers what he takes to be the most developed, plausible ESC account, Daniel Hutto’s Narrative Practice Hypothesis (NPH). NPH is a non-mindreading account of how children develop folk psychology.² In addition to our primary embodied practices, as adults, we also rely on folk psychology as a more sophisticated way to understand others’ behaviors. NPH aims to explain how children develop this ability to understand behavior in terms of *reasons* without attributing to them implausibly precocious mindreading abilities. NPH holds that the source of our capacity for sophisticated social cognition is direct encounters with folk psychological narratives, stories that exemplify the forms and norms of social interactions. Stories like Little Red Riding Hood and Goldilocks and the Three Bears are paradigmatic folk psychological narratives. The narratives provide exemplars of how agents act according to reasons in order to attain some goal. The child and her caretaker jointly attend to the narrative, and through guided interaction with a caretaker, the child becomes acquainted with forms and norms of folk psychology. On this view, developing folk psychological competence consists in learning how to give and receive reason-giving explanations. NPH is meant to be completely independent from mindreading. It does not consist in, nor does it depend on, mindreading (Gallagher and Hutto 2007; Hutto 2008).

NPH holds that to be proficient in folk psychology, children must first come to have and attribute propositional attitudes, they must learn how propositional attitudes—e.g., beliefs, desires, and emotions—combine to form reasons for

¹ For a critical assessment of these claims, see Spaulding (2010).

² Although “folk psychology” and “mindreading” are sometimes used interchangeably, following Hutto, I shall not use the terms this way. Folk psychology is the practice of understanding others’ behavior in terms of reasons for actions, whereas mindreading is the practice of ascribing beliefs and desires to a target in order to explain and predict the target’s behavior. These are distinct notions. Folk psychological competence may or may not involve a capacity for mindreading. Hutto argues that it does not.

action, and, through exposure to folk psychological narratives, they learn the norms of folk psychology. Only after we master folk psychology can we learn to mindread—to *explain* and *predict* behavior on the basis of attributed mental states. But even then, mindreading is a rarely used, specialized skill. We mindread only when our primary modes of social cognition break down.

Because NPH holds that we develop our folk psychological skills by comprehending folk psychological narratives, the view implies that children do not understand folk psychology until after they develop language skills. There is, thus, a chasm between the preverbal social cognition of infants and that of older children and adults who have mastered language. On this view, only those who have mastered language and encountered folk psychological narratives can have folk psychological competence. Nonlinguistic creatures cannot be *real* folk psychologists because their understanding of others is limited to nonpropositional, nonrepresentational, and nonmental understanding. They are capable of engaging in embodied intersubjective practices, but to understand how others act for reasons much more is required.

Thompson argues against the developmental timeline proposed by NPH. Recent findings in developmental psychology provide evidence against the idea that there is a deep chasm between preverbal and verbal children's social cognition and that preverbal children are limited to nonpropositional, nonrepresentational, nonmental understanding of others. It is worth reviewing the evidence from developmental psychology because several articles in this special issue discuss these findings.

For much of the last 30 years, the standard developmental picture of mindreading has been that at around 4 years of age, children undergo a fundamental shift in their mindreading abilities. As Heinz Wimmer and Josef Perner's experiments first revealed, and other experiments have since replicated, before the age of 4, children cannot pass standard false-belief tasks (Wimmer and Perner 1983; Gopnik and Astington 1988). In one task commonly referred to as the *Sally-Anne task*, children listen to a story as it is enacted with dolls named Sally and Anne. In the scene, Sally hides a toy in one place and then she leaves the scene. Anne moves the toy from the original hiding place to a new hiding place. When children younger than 4 years old are asked where Sally will look for the toy, they answer incorrectly. They say she will look in the new hiding place. Children 4-years and older, however, typically answer correctly. They say Sally will look in the original place and give appropriate explanations for why she will look there. This evidence has been taken to show that there is a significant developmental shift in mindreading abilities at around 4 years of age. At age 4, children shift from lacking proficiency with the concept of *belief* to being able to appropriately apply the concept in a range of situations. That is, at age 4, children master the *belief* concept. Given that the concept of *belief* plays an important role in understanding others' mental states, the standard false-belief task has been taken to be the measuring stick of mindreading abilities.

Onishi and Baillargeon (2005) have objected to the standard false-belief tasks, arguing that these tasks are computationally and linguistically too taxing for children younger than 4 years old. The standard false-belief task requires children to remember the details of the story, who saw what and when, to interpret adults' questions and give appropriate responses to these questions. Many of these task demands are unrelated to mindreading per se. Rather, the demands of the standard false-belief task reveal performance of executive functions, e.g., memory and response inhibition.

In lieu of the standard measuring stick, Onishi and Baillargeon opt for a simplified nonlinguistic false-belief task to measure mindreading abilities of younger children.

In their novel nonlinguistic false-belief task, 15-month-old infants watch an actor put a toy watermelon slice in one of two adjacent boxes, a green box or yellow box. Next, the toy is moved. In half the trials, the toy is moved halfway to the other box and then back to the original box, and in the other half of the trials, the toy is moved to the other box. For both of these conditions, the actor either does or does not see the movement of the toy. (In one variation, she looks through an opening in the tops of the boxes, and in another variation, she does not.) Using the violation-of-expectation method, Onishi and Baillargeon found that 15-month-old infants looked longer in two cases: first, when the actor does not see that the toy's location has changed but searches in the correct box anyway and, second, when the actor does see the toy being relocated but the actor reaches in the incorrect box.

Onishi and Baillargeon interpret these results as showing that the 15-month-old infants expect the actor to search for the toy on the basis of her belief about the toy's location. When the actor does not search for the toy on the basis of her belief, the infants' expectations are violated, and they thus looked longer at those events. Onishi and Baillargeon take this to be good evidence for the conclusion that 15-month-old infants already have mindreading abilities and that the ability to mindread, in at least a rudimentary form, is innate. Employing a variety of nonverbal testing methods—anticipatory looking, violation of expectation, and active helping—several other theorists have found evidence that preverbal infants are sensitive to others' mental states, e.g., others' preferences and beliefs about the identity, properties, and location of objects.

Thompson argues that these findings are troublesome for ESC. These data suggest that preverbal infants ascribe mental states to others, which, if true, undermines the idea that preverbal children are limited to nonpropositional, nonrepresentational, nonmental understanding of others. If Thompson's interpretation of these studies is correct, then non-mentalistic primary and secondary intersubjectivity cannot account for the social cognition of preverbal infants, and NPH seems to have things exactly backwards. It is not the case that participation in narrative practice allows us to develop folk psychological and mindreading abilities. To the contrary, Thompson argues that in order to participate in narrative practices, one must already possess mindreading abilities. Narrative practice presupposes mindreading abilities. That is, we understand folk psychological narratives because we have mindreading abilities. Thompson concludes that the best explanation of all the data is this cognitivist explanation.

Herschbach

In the second article criticizing ESC, Mitchell Herschbach considers a more radical version of ESC: enactivism. Enactivism holds that social interaction *constitutes* social cognition. This is not merely a methodological claim that in studying social cognition, we should study interactive practices. As Herschbach explains, the enactivist claim is much stronger than that. Enactivism holds that the interaction between two autonomous agents can become coupled in a way such that the interaction becomes autonomous itself and this autonomous process *constitutes* social cognition.

Enactivists Hanne De Jaegher and Ezequiel Di Paolo claim that “Social interaction is the regulated coupling between at least two autonomous agents, where the regulation is aimed at aspects of the coupling itself so that it constitutes an emergent autonomous organization in the domain of relational dynamics, without destroying in the process the autonomy of the agents involved (though the latter’s scope can be augmented or reduced)” (De Jaegher and Di Paolo 2007, p. 493). Enactivists argue that “social cognition is not reducible to the workings of individual cognitive mechanisms” and that “interactive processes are more than a context for social cognition: they can complement and even replace individual mechanisms” (De Jaegher et al. 2010, p. 441). The co-regulated coupling between two autonomous agents itself is social cognition, and it is not reducible to one or the other agent’s cognitive processes or internal neural mechanisms, e.g., a mindreading or theory of mind mechanism.

Herschbach argues that enactivism suffers from conceptual confusions regarding what it means for interaction to *constitute* social cognition. De Jaegher et al. (2010) distinguish the following three roles that interaction might play in explaining social cognition: *enabling condition*, *contextual factor*, and *constitutive element*. Although this taxonomy is introduced to avoid conceptual confusion, there is an ambiguity in the way enactivists treat constitutive elements. Enactivism relies on two distinct senses of being a constituent in a cognitive action: a *compositional* sense of constitution (an element E is part of action A) and a *causal* sense of constitution (E helps produce A). Consideration of the examples enactivists use in support of their account shows that in cases where social interaction is necessary for social cognition, sometimes enactivists treat the social interaction as a constitutive element, whereas at other times they treated it as a mere enabling condition. This equivocation makes a muck of enactivists’ conclusions about social cognition

For example, if the compositional sense is the appropriate notion of constitution, then this view is inconsistent with some of the standard enactivist conclusions (e.g., that social interaction in perceptual crossing experiments constitutes social cognition). If, however, the causal sense is the best way to understand constitution, then this view is inconsistent with other enactivist conclusions (e.g., that protoconversations—face-to-face interactions between parents and infants—merely enable social cognition). This ambiguity pervades enactivist writings. Herschbach concludes that this taxonomy is inadequate for distinguishing constitutive elements of a cognitive system from external elements that merely causally interact with it.

Herschbach offers a mechanistic account as an alternative to enactivism. Mechanism, he argues, can adequately explain the cases enactivists aim to explain without the resulting conceptual confusions. Mechanistic explanations involve decomposing a system into its parts and determining what operations those parts perform. The mechanistic account Herschbach offers is compatible with cognitivist explanations of social cognition, such as explaining mental state attribution in terms of internal neural mechanisms whose operations can be described in terms of mindreading. Though mechanistic accounts are compatible with mindreading accounts, they also encourage *looking around* to the environment in which the mechanism is situated. For example, *dynamic mechanistic explanation* can make use of the tools of dynamical systems theory to characterize the temporal organization of parts and operations in a mechanism and the mechanism’s activity in its environment. Dynamic mechanistic explanation, Herschbach argues, captures

much of the motivation for enactivism—and arguably other ESC accounts. That is, it emphasizes an organism’s situatedness, or embeddedness in an environment.

Herschbach applies the mechanistic account to the enactivists’ examples and finds that the mechanistic account can better handle these cases without the resulting conceptual confusion that plagues enactivism. The mechanistic account emphasizes the explanatory roles of elements at multiple levels of organization, e.g., the social, individual/personal, and subpersonal levels, while distinguishing the constituents of the individual agent from environmental factors influencing the agent. Thus, Herschbach offers an alternative to enactivism that is compatible with cognitivist accounts of social cognition and is capable of capturing the motivations of enactivism.

Reconciliation Accounts

The previous two articles mainly focus on criticizing existing ESC accounts. The next four articles take as starting points ideas the authors regard as genuine insights of ESC. Each paper attempts to modify, temper, or excise problematic elements of ESC. In doing so, these papers construct ESC-inspired accounts that are less averse to cognitivism.

Zawidzki

In his article “Unlikely allies: Embodied social cognition and the intentional stance,” Tadeusz Zawidzki articulates a middle-ground position between ESC and cognitivism. Zawidzki explains that ESC accounts are a response to *metapsychological* theories of social cognition, e.g., TT and ST, which dominated the literature for the last 20 years. ESC theorists find the presumption that social cognition is based on mental state attributions a fundamentally flawed idea. Typically, ESC theorists reject the following three ideas: (a) metapsychology, (b) cognitivism, and (c) skepticism about the relevance of phenomenology in an account of the cognition. In addition to rejecting metapsychology, proponents of ESC explicitly criticize cognitivism as it is the basis of metapsychology, and they are sanguine about using phenomenological insights to justify claims about how human beings navigate their social worlds.

Zawidzki argues that there is no need for ESC theorists to reject (b) and (c) in constructing an alternative to (a). He says, “Despite the close connections that ESC theorists see between skepticism about metapsychology, the critique of cognitivism, and the embrace of phenomenology, these assumptions do not mutually imply each other”. The motivating idea behind ESC is a rejection of metapsychology, but rejecting metapsychology does not require hostility toward cognitivism, nor does it require reliance on phenomenology. In fact, Zawidzki argues, the ESC arguments rejecting cognitivism and endorsing phenomenological arguments about the mind are problematic.

Zawidzki offers a middle-ground account of social cognition based on Daniel Dennett’s Intentional Stance (IS). Dennett’s IS holds that to interpret a system as intentional is to assume that its behavior is a rational response to its circumstances

given its goals. Taking IS as a model for quotidian interpretation of others' behavior, understanding others involves tracking abstract patterns, not attributing unobservable, concrete mental causes. Although Dennett characterizes IS in terms of belief and desire ascriptions, Dennett holds a heterodox view of propositional attitudes. Beliefs and desires, for Dennett, are abstract, instrumentalist posits more akin to "relevant information" and "relevant goal" than to "concrete, unobservable mental states with representational content and causal influence over behavior."

Zawidzki argues that not only is this the correct interpretation of IS, but this interpretation can also be used to resolve apparent problems with IS. For example, many have criticized IS's instrumentalist analysis of propositional attitudes. However, if we understand IS as offering an analysis of *quotidian* interpretation, one can accept criticisms of IS as an account of the propositional attitudes while still endorsing it as a model of quotidian interpretation. Furthermore, recent empirical evidence from developmental psychology supports this account. Gergely Csibra and Gyorgy Gergely found that infants as young as 6 months interpret others' behavior as rational. Infants naturally adopt a so-called *teleological stance*. Agents are expected to perform the most efficient means–ends action available to them within their situational constraints to bring about the goal state. Thus, Csibra and Gergely's teleological stance nicely complements IS.

IS and ESC are alike in rejecting metapsychological accounts of social cognition. Zawidzki argues that an alliance between IS and ESC would be mutually beneficial. On the one hand, this interpretation of IS can be used to explicate various underdeveloped ESC concepts, such as primary and secondary intersubjectivity and direct social perception. IS can characterize *how* neural processes bring information to bear on direct social perception. On the other hand, ESC can provide phenomenological insights and illuminate the role of contexts and interaction in direct social perception that IS misses.

This alliance between IS and ESC provides a way to save the motivation for ESC—dissatisfaction with metapsychology—without committing ESC to implausible phenomenological arguments and unnecessary resistance toward cognitivism. Moreover, the alliance strengthens ESC by demystifying some of the crucial concepts of ESC. Thus, Zawidzki offers an account that occupies a middle ground between cognitivism and standard ESC accounts.

Shepherd

Joshua Shepherd in "Action, mindreading and embodied social cognition" offers the second reconciliation type of account. Shepherd highlights the fact that a central guiding principle of embodied cognition is that cognition is closely tied to action, and social cognition is no exception. ESC's main complaint against mindreading accounts is that they portray social cognition as a spectator sport, whereby we understand others by taking a passive third-person observational stance from which we explain and predict others' behaviors. ESC is motivated by the idea that social cognition is an inherently interactive embodied practice.

Although hard-line ESC and mindreading theorists may view these accounts as competing paradigms, Shepherd does not view the debate this way. Shepherd argues that a profitable way of understanding ESC is as a source of insight about the process

of mindreading. This account is a departure from ESC in two ways. First, ESC theorists typically argue that mindreading is a rarely used, specialized skill, not something fundamental to social cognition. Shepherd rejects such demotion claims about mindreading. He argues that mindreading is one—though not the only—important aspect of social cognition. Second, ESC theorists typically accept mindreading theorists' characterization of the process of mindreading. Shepherd argues that we ought to question the typical characterization of mindreading processes. It is here that he sees a role for action-based cognition that is often emphasized by embodied cognition. *Actional processes and states* (APS) consist in an agent's implicit or explicit goals, intentions, desires, emotions, needs, beliefs about actional capacities, etc. Shepherd's project is to motivate a study of the role of APS in mindreading by showing how plausible it is that they play some kind of role in mindreading. Understanding the roles that APS play in mindreading is important to the overall aim of understanding mindreading processes and their role in cognition more broadly.

Shepherd argues that it is plausible that APS play important roles in mindreading. He cites a wealth of data that support the idea that APS influence our cognitive processes in general. Applying this general lesson to social cognition, Shepherd lays out empirical evidence suggesting that representations of one's own capacities for action influence one's social perception. He argues that APS constrain operations of mindreading processes in three ways: (1) by motivating agents' engagement in these processes in certain ways, (2) making certain information salient, and (3) directing attention in social contexts. Below, I shall canvas some of the data Shepherd cites in support of his thesis.

In one study, participants were instructed to visualize and write in detail about a time in their lives when they were sexually and romantically aroused. Primed with a mate-search motive, participants viewed a series of faces of both attractive and unattractive members of both sexes. Participants were then tested for their level of socio-sexual orientation and classified as sexually restricted or sexually unrestricted. Sexually unrestricted participants who received the mate-search prime took longer to disengage their attention from attractive opposite-sex faces, suggesting that an individual goal influenced the direction of attention to socio-cognitive inputs (Maner et al. 2007a, b). Moreover, when primed with a mate-search motive, males were more likely to attribute sexual arousal to attractive female faces but not to average-looking female faces (Maner et al. 2005).

Shepherd also cites data indicating that experiencing social exclusion heightens motivation for connection with others and influences the judgments participants make regarding others. Researchers found that subjects who were primed for social exclusion tended to rate the target individuals as nicer, friendlier, and more desirable than those primed for social inclusion. This suggests that being socially excluded leads one to have an implicit goal to reconnect with others, and this goal influences one's social judgments in a specific direction (Maner et al. 2007a, b). These and other data highlight the relevance of APS in social cognition.

Shepherd's account takes as its departure point the ESC claim that social cognition is fundamentally action and interaction based. Given the empirical data suggesting that APS play an important role in social cognition, ESC accounts are right to emphasize APS. However, Shepherd notes, mindreading accounts of social cognition

seem to leave no room for APS. They are by and large silent on the issue. Thus, Shepherd's goal is to argue that any adequate account of mindreading ought to take into consideration APS. His account uses this genuine insight of ESC to understand mindreading processes and their role in cognition more generally. Thus, this account aims to reconcile ESC accounts of social cognition and cognitivist accounts of mindreading.

Wilby

In his article, "Embodying the false belief tasks," Michael John Wilby notes that ESC accounts often define themselves in contrast to cognitivist accounts. Like other authors in this special issue, Wilby argues that ESC theorists make useful insights regarding the limits of cognitive psychology in the study of social cognition. However, ESC's tendency to focus primarily on the relation between the phenomenology of mindreading and the neuroscience of social cognition results in neglect of other important empirical research. Specifically, ESC theorists often dismiss findings in developmental psychology as tainted with theoretical presuppositions from cognitivism. The result of such a dismissal is that theorists run the risk of ignoring important psychological data and not explaining the connection between lower-level social cognition and higher-level social cognition.

The standard false-belief task, as discussed in Section 1.1, is regarded in contemporary literature as an idiosyncratic task that tests a *variety* of skills, e.g., the ability to represent a target's mental states, inhibit prepotent responses, and interpret and respond correctly to linguistic commands and questions. Given that passing the standard false-belief task requires these other non-mindreading skills, the task is no longer regarded as a good measuring stick for the development of mindreading. Nevertheless, passing the standard false-belief task is a developmental milestone. It is a robust finding, having been replicated many times, that children do not pass the false-belief task until they are about 4 years old.

As I discussed in Section 1.1, recent findings in developmental psychology suggest that preverbal 15-month-old infants are sensitive to the false beliefs of others. This is puzzling. These new wave nonlinguistic false-belief tasks reveal that infants just over a year old are sensitive to others' false beliefs. Children become adept language users, capable of linguistically expressing their own beliefs and desires, in the second year of life. But it is not until age 4 that they can linguistically express their capacity to recognize false beliefs. Why is there such a significant time lag between passing the nonlinguistic false-belief task and the standard false-belief task? This is what Wilby calls the *time lag problem*, and he argues that any adequate account of social cognition needs to solve the time lag problem.

Developmental evidence of time lag and the resulting problem of understanding social cognition before and after 4 show the need for a comprehensive account of social cognition. Wilby argues that ESC on its own cannot offer a comprehensive account of social cognition. He argues that we need something to bridge the gap between the basic embodied practices that ESC emphasizes and the metapsychological skills, e.g., false-belief reasoning that mindreading accounts emphasize. Wilby considers and rejects one prominent bridging account, Ian Apperly and Stephen Butterfill's Minimal Theory of Mind account (2009). Apperly and Butterfill posit a

two-systems account of social cognition according to which system 1 is fast, inflexible, limited in the kind of content it can process, and involved in predicting actions based on behavioral regularities, whereas system 2 is slower, flexible, virtually unlimited in the kind of content it can process, and involved in explaining and predicting others' behavior in terms of mental states. Apperly and Butterfill—and other proponents of two-systems accounts of social cognition—argue that infants are capable of only system 1 social cognition whereas older children are capable of system 2 social cognition.³

Wilby argues that although this account is on the right track insofar as it aims to explain social cognition before and after age 4, it fails to solve the time lag problem. There remains an unbridgeable gap between system 1 social cognition and system 2 social cognition. That is, Apperly and Butterfill's account does not explain *how* children develop mature mindreading skills. Apperly and Butterfill do not tell us *how* children develop system 2 cognition from system 1 cognition. Moreover, it is not clear that their account could tell us how this happens.

Wilby offers the notion of a *joint mental state* to bridge the gap between basic embodied practices and mindreading. A joint mental state is “one in which the understanding that each have of the other is not of the form <X perceives that <Y perceives that p>> and <Y perceives that <X perceives that p>>, but rather of the form <X and Y jointly perceive that p>. The participants' relation to each other is as cosubjects, not as objects of each other's attention”. Neither X nor Y need token the thought, “We jointly perceive that p.” Rather more radically, only the joint system itself tokens a thought with the content “p.” The notion of a joint mental state is in the spirit of the enactivist account of social cognition insofar as the joint mental state is attributed to a single dynamical system constituted by two interacting individuals. (See Section 1.2 for discussion of enactivist social cognition.) Moreover, the joint mental state provides an ESC-inspired bridge between lower-level social cognition and higher-level social cognition.

Contrasting his view with Apperly and Butterfill's account, Wilby argues that his view provides an entry point into the mature understanding of others, whereas Apperly and Butterfill's account does not. On Wilby's view, when the infant is in a joint mental state with her caretaker, the infant stands in a psychological relation to the other agent. This psychological relation is necessary (though not sufficient) to bridge the gap between nonpsychological understanding of others' behavior and the more sophisticated psychological understanding of others.

Joint mental states can be used to explain the new wave false-belief tasks and the development of a sophisticated mode of social cognition. They can be used to connect the basic embodied interactive practices (i.e., primary intersubjectivity) that ESC emphasizes with the metapsychological understanding of others that cognitivist accounts emphasize. Wilby argues that in order for ESC to offer a comprehensive picture of social cognition, it must focus on an often-neglected part of the empirical literature—developmental psychology—to explain how children develop

³ Thus, the two-systems accounts offer a deflationary explanation of the new wave false-belief tasks. The nonlinguistic false-belief tasks that 15-month-olds pass do not indicate that infants are tracking *false beliefs*, according to this view. Instead, these infants are tracking what Apperly and Butterfill call “belief-like states.” See Wilby's article in this issue and Apperly and Butterfill (2009) for more on exactly what “belief-like states” are.

sophisticated metapsychological skills. Doing this requires less hostility toward cognitivism and more cooperation.

de Bruin and Kästner

In “Dynamic embodied cognition,” Leon de Bruin and Lena Kästner argue for an account of social cognition that mediates between cognitivism and enactivism. Like the other three papers in this section, this paper argues that neither the ESC view nor the traditional cognitivist view presents a compelling account of social cognition. Enactivism, de Bruin and Kästner argue, too heavily emphasizes the role of online (“coupled”) processes in social cognition and neglects the role of offline (“decoupled”) processes. Cognitivist accounts do just the opposite: they emphasize offline processes at the expense of online processes.

Enactivism holds that cognition is a process of sense making that emerges from the dynamic online interaction or “coupling” between autonomous agents and the environment in which they are embedded. Enactivism focuses on explaining cases of social cognition that are carried out online, e.g., gaze following, joint attention, imitation. Given the prominence of online processes in this account, enactivists tend to ignore the significant role that *offline* processes play in cognition. Offline processes involve internal representations, which are not bound to the current features of the agent’s body or her environment and hence “decoupled” from the agent’s environment, e.g., first-order and second-order belief ascriptions. It is undeniable that we can and sometimes do engage in offline processes. Enactivists often simply ignore such offline processes. Thus, enactivism leaves a “cognitive gap” between online and offline social cognitive processes. Furthermore, appeals to phenomenology merely obfuscate this matter.

There is a parallel problem with cognitivism. Cognitivism holds that the mind is an intracranial information-processing system that manipulates symbolic representations. On this view, cognition just is this computational process. Cognitivist accounts of social cognition tend to focus on our abilities to explain and predict others’ behaviors in terms of mental state attributions. They explain these social cognitive capacities by adverting to computational processes, modular mechanisms, internal representations, and so on. Given the emphasis on these offline computational processes, cognitivism tends to neglect the role of *online* processes in social cognition leaving the same kind of cognitive gap as enactivism.

De Bruin and Kästner’s own account, Dynamic Embodied Cognition (DEC) starts from an embodied view of cognition. It not only emphasizes exploratory interaction with one’s environment, like most versions of enactivism, but also includes a significant role for offline processes in social cognition. Offline processes typically involve the manipulation of information that is absent from the environment and has to be internally represented. Taking processes offline can be advantageous because it allows the agent to withdraw from the immediate surroundings so as not to automatically act upon particular affordances. Think of the enormous cognitive benefit the capacity to engage in counterfactual reasoning affords. Instead of simply engaging directly with the environment, the agent may consider other ways to respond to the environment, consider hypothetical action plans, evaluate various means of achieving one’s goals, etc. An agent’s online and offline processes interact to yield cognitive

flexibility and autonomy from environmental stimulation such that the agent becomes less dependent upon, and gains new ways of relating to, her environment and other agents in her environment. De Bruin and Kästner argue that DEC, unlike enactivism and cognitivism, plausibly explains the role of both online and offline processes and the ways in which they interact.

DEC can be modeled using Dynamic Systems Theory. This provides a way of understanding online and offline processes in a unified way. Moreover, DEC can be used to explain different results in developmental psychology. Specifically, de Bruin and Kästner suggest that we understand the cognitive gap (or time lag, as Wilby calls it) between passing the nonlinguistic false-belief tasks and the standard false-belief task in terms of the degrees of decoupling required. In order for one's expectations to be violated in the nonlinguistic false-belief task, one must be capable of a minimal amount of decoupling from the environment. Even 15-month-old infants are capable of decoupling from their environment, as their success on nonlinguistic false-belief tasks demonstrates. From 15 months to 4 years, children gain the capacity to decouple from their environment to a greater degree. Thus, according to DEC, the difference between the new wave false-belief tasks and the standard false-belief task is the degree to which they require offline processing. And because cognitive development is a matter of degree (codifiable by Dynamic Systems Theory) not of kind, DEC does not suffer from an insurmountable cognitive gap like enactivism and cognitivism. DEC aims to capture the insights of both enactivism and cognitivism, avoid the pitfalls of these views, and offer a comprehensive account of social cognition.

Applications of ESC

The articles discussed in the first section primarily criticize ESC, whereas the articles discussed in the second section identify problems with ESC and attempt to construct accounts that are more conciliatory toward cognitivism so as to avoid those problems. The focus of the articles in Section 3 is somewhat different. The three articles I will discuss in this section apply ESC to particular domains of cognitive science, specifically, mirror neurons, meaning and linguistic content, and gestures. Although the authors of these articles sometimes offer criticism of extant ESC theories, their papers can be seen as individually and collectively making the case for ESC by showing how it can be applied fruitfully to specific cases.

Michael

In "Mirror systems and simulation: A neo-empiricist interpretation," John Michael argues that we can use an embodied cognition-inspired concept of *simulation* to understand the cognitive function of mirror neurons in social cognition. Mirror neurons are neurons that fire or activate when a subject acts or emotes and also when a subject observes a target acting or emoting. For example, a host of neurons in the premotor cortex and posterior parietal cortex fires when I grasp an object, and this

same host of neurons fires when I observe another person grasping an object. There are similar neural systems in the amygdala and insula for experiencing and observing disgust and fear (Rizzolatti and Craighero 2004). For the sake of simplicity, I shall focus only on motor mirror neurons. In execution mode, mirror neuron activation in the premotor and posterior parietal cortex constitutes a motor intention. In observation mode, these very same neurons activate. Many theorists argue that mirror neuron activation in observation mode also constitutes an intention—one that is tagged as belonging to the target. Our mirror neurons activate as if we are acting as the target is acting. Thus, it seems that the process of understanding a target's behavior involves simulating the target's motor intentions.

These data indicate that mirror neurons involve simulation. As a result, many have argued that mirror neurons are evidence for the ST of mindreading.⁴ Michael points out, however, that there are a variety of different models of simulation involved in mirroring, not all of which are relevant to ST. He identifies four models of simulation: *direct matching*, *inverse modeling*, *predictive coding*, and *response modeling*. Each model of simulation can be used to describe some aspect of a mirror neurons' functional role. Michael argues for a broader conception of simulation that can encompass each of these models of simulation.

The concept of simulation Michael advocates is an extension of the neo-empiricist theory of concepts (Barsalou 1999; Lakoff and Johnson 1999). Neo-empiricism is the view that conceptual thought in general involves simulation in the sense that it is grounded in sensory, motor, and other embodied systems. Learning a concept involves reinforced activation of particular sensory and motor representations, and employing a concept in a cognitive task involves reactivating these sensory and motor representations. Michael argues that the neo-empiricist account can subsume the various conceptions of simulation, thereby offering a framework in which each of these models of simulation may play a role. The neo-empiricist super model can frame the ways in which mirror neurons contribute to social cognition in a more unified and plausible way than any of the individual models of simulation. Moreover, the neo-empiricist model is capable of explaining the ways in which the more specific models of simulation relate to each other. Hence, applying the neo-empiricist ESC account to mirror neurons yields a fruitful way of understanding their role in social cognition.

Sandler

In "Reenactment: An embodied cognition approach to meaning and linguistic content," Sergiy Sandler argues that we can apply ESC to a surprising domain: meaning and linguistic content. Sandler argues that *reenactment*, a concept similar to Michael's neo-empiricism, can be used to explicate different aspects of linguistic meaning and forms of utterances. Reenactment, as Sandler uses the term, is the overt embodied simulation of action. Recent empirical data on mirror neurons suggest the process of understanding others' actions involves reenactment. That is, understanding an action involves

⁴ Although, see Spaulding ([forthcoming](#)) for an alternative account of the cognitive function of mirror neurons in social cognition and Spaulding (2012) for an argument that mirror neurons are not evidence for ST.

executing some of the processes involved in those actions. (See Section 3.2 for more on mirror neurons.) Sandler argues that understanding others' linguistic meaning also involves reenactment. His argument takes the form of a plausibility argument: he aims to show that reenactment linguistics is plausible by showing how established theories can be stated in reenactment terms. Sandler does not advocate replacing existing theories. Rather, his project is to show how the conceptual framework of reenactment is useful.

Sandler considers several different aspects of linguistic meaning and forms of organizing utterances and shows how they can be described as forms of reenactment. Specifically, he discusses (1) reference to remote or concealed entities in discourse, (2) the use of grammatical constructions, (3) complex enunciational structures, (4) the segmentation of long utterances, and (5) the overall pragmatic meaning of utterances. The details are different for each case, but the unified theme is that understanding others' linguistic utterances plausibly involves embodied reenactment. Sandler's project bears much in common with Gallese and Lakoff's (2005) attempt to ground the semantics of *concepts* in embodiment and Lawrence Barsalou's perceptual symbol systems. The ideas articulated in Sandler's article relate to embodied approaches to social cognition in that they ground one important kind of social interaction—linguistic communication—in embodiment.

Cuffari

In “Gestural sense-making: Hand gestures as intersubjective linguistic enactments,” Elena Clare Cuffari argues ESC offers a fruitful way of understanding the role of gestures in cognition. Cuffari argues that spontaneous co-speech gesturing, the ubiquitous human practice of spontaneously gesturing while speaking, confirms embodied aspects of linguistic meaning making.

Cuffari considers a *meaning-leaking* paradigm, which regards spontaneous co-speech gesturing as merely uncontrollable, unconscious windows into speakers' thoughts. The meaning-leaking paradigm draws a sharp boundary between language, which is regarded as normative and conventional, and gestures, which are regarded as spontaneous, nonnormative, and nonconventional. Cuffari argues the meaning-leaking paradigm oversimplifies the phenomena and overlooks the irreducible sociality of linguistic activity (broadly construed to include co-speech gestures). Co-speech hand gestures are first and foremost emergent elements of social interaction, not merely epiphenomenal products of an isolated internal computational process.

Cuffari argues instead for a *meaning-building* paradigm, according to which gestures are embodied performances of intentions that are accessible to both speakers and listeners. Gestures are constitutive elements of communicative dialogue. The speaker and listener “co-author” the emerging dialogue. Moreover, it is precisely this intersubjectivity of dialogue that makes linguistic activity (including gestures) an inherently normative dimension of communicative action. Verbal language and embodied communication, e.g., visible kinesthetic bodily practices, are evaluated not in terms of competence, but in terms of the intersubjective, pragmatic communicative standards that communities build together. Cuffari's

project of grounding gestures in embodied, meaningful intersubjective interaction is complementary to Sandler's account of reenactment linguistics.

Assessment

I have highlighted three structural themes of the papers in this special issue. The papers discussed in Section 1 primarily criticize ESC; the papers discussed in Section 2 construct ESC-inspired accounts that reconcile cognitivism and ESC; and the papers discussed in Section 3 highlight the merits of ESC by showing how it can be fruitfully applied to particular socio-cognitive phenomena. Beyond these structural themes, there are commonalities amongst the papers in their critiques and praise of ESC.

In this collection of papers, the most common criticism of ESC is that, as it is standardly presented, it is not a comprehensive account of social cognition. Thompson argues that ESC cannot account for preverbal infants' precocious mindreading skills that are demonstrated by the new wave false-belief tasks. Given that ESC accounts regard preverbal children as limited to non-mentalistic understanding, ESC cannot explain the data suggesting that preverbal infants are sensitive to others' false beliefs. Wilby argues that ESC cannot solve the time lag problem. That is, it cannot explain how children develop sophisticated socio-cognitive skills from basic embodied practices. Shepherd argues that ESC accounts mistakenly demote the status of mindreading in social cognition, thereby neglecting an important element of social cognition. And de Bruin and Kästner argue that ESC's exclusive focus on online "coupled" processes leaves it unable to explain the cognitive role of offline processes and how these interact with online processes. Hence, in various ways, these authors object that ESC is not a comprehensive account of social cognition.

However, the authors in this special also highlight the genuine insights of ESC. Although Herschbach ultimately rejects enactivism, his own positive account, *dynamic mechanistic explanation*, involves an emphasis on the situatedness or embeddedness of a mechanism in an environment. Herschbach's account uses the tools of dynamical systems theory to characterize the temporal organization of parts and operations in a mechanism and the mechanism's activity in its environment. Zawidzki argues that ESC is right to be skeptical of metapsychology as an account of social cognition. Shepherd argues that ESC's emphasis on actional processes can shed light on the process of mindreading. De Bruin and Kästner argue that ESC is right to bring attention to online socio-cognitive processes such as gaze following and joint attention. Such processes are fundamental to social cognition and often ignored by cognitivist accounts of social cognition. Finally, the last three articles demonstrate the insights of ESC by applying it to particular domains. If these analyses are successful, this is good reason to think that ESC is a useful, fruitful paradigm for understanding social cognition.

The debate between ESC and cognitivism is, of course, not over. Many disputes between ESC and cognitivism remain unresolved. The aim of this special issue is not to come to a verdict about ESC, but rather to assess the current state of the debate between ESC and cognitivism. The goal is to highlight genuine insights of ESC and critically evaluate controversial elements with the hope of bringing clarity and progress to the study of social cognition.

References

- Apperly, I., & Butterfill, S. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*, 953–970.
- Barsalou, L. W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences*, *22*, 577–609.
- Bermúdez, J. L. (2003). The domain of folk psychology. In A. O’Hear (Ed.), *Minds and persons* (pp. 25–48). Cambridge: Cambridge University Press pp.
- Carruthers, P., & Smith, P. K. (1996). *Theories of theories of mind*. Cambridge: Cambridge University Press.
- Currie, G., & Ravenscroft, I. (2002). *Recreative minds: imagination in philosophy and psychology*. Oxford: Oxford University Press.
- Davies, M., & Stone, T. (1995a). *Folk psychology: the theory of mind debate*. Oxford: Blackwell.
- Davies, M., & Stone, T. (1995b). *Mental simulation: evaluations and applications*. Oxford: Blackwell.
- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences*, *6*(4), 485–507.
- De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, *14*(10), 441–447.
- Gallagher, S. (2005). *How the body shapes the mind*. USA: Oxford University Press.
- Gallagher, S. (2009). Neural simulation and social cognition. In J. A. Pineda (Ed.), *Mirror neuron systems* (pp. 1–17). New York: Humana.
- Gallagher, S., & Hutto, D. D. (2007). Understanding others through primary interaction and narrative practice. In Sinha, Itkonen, Zlatev, & Racine (Eds.), *The shared mind: perspectives on intersubjectivity*. Amsterdam: John Benjamins.
- Gallese, V., & Lakoff, G. (2005). The brain’s concepts: the role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*, 455–479.
- Goldman, A. I. (2006). *Simulating minds: the philosophy, psychology, and neuroscience of mindreading (philosophy of mind)*. USA: Oxford University Press.
- Gopnik, A., & Astington, J. W. (1988). Children’s understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*(1), 26–37.
- 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.
- Gordon, R. M. (1986). Folk psychology as simulation. *Mind and Language*, *1*, 158–171.
- Hutto, D. D. (2004). The limits of spectatorial folk psychology. *Mind & Language*, *19*(5), 548–573.
- Hutto, D. D. (2008). *Folk psychological narratives: the sociocultural basis of understanding reasons*. Cambridge: MIT Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: the embodied mind and its challenge to western thought*. New York: Basic Books.
- Maner, J., Kenrick, D., Neuberg, S., Becker, D. V., Robertson, T., Hofer, B., Delton, B., Butner, J., & Schaller, M. (2005). Functional projection: how fundamental social motives can bias interpersonal perception. *Journal of Personality and Social Psychology*, *88*, 63–78.
- Maner, J., DeWall, N., Baumeister, R., & Schaller, M. (2007a). Does social exclusion motivate interpersonal reconnection? Resolving the “porcupine problem”. *Journal of Personality and Social Psychology*, *92*, 42–55.
- Maner, J., Galliot, M., Rouby, D. A., & Miller, S. (2007b). Can’t take my eyes off you: attentional adhesion to mates and rivals. *Journal of Personality and Social Psychology*, *93*, 389–401.
- Nichols, S., & Stich, S. P. (2003). *Mindreading: an integrated account of pretence, self-awareness, and understanding other minds*. Oxford: Oxford University Press.
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255–258.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*(1), 169–192.
- Shapiro, L. S. (2010). *Embodied cognition*. Abingdon: Routledge.
- Spaulding, S. (2010). Embodied cognition and mindreading. *Mind and Language*, *25*(1), 119–140.
- Spaulding, S. (2012). Mirror neurons are not evidence for the simulation theory. *Synthese*. doi:10.1007/s11229-012-0086-y.
- Spaulding, S. (forthcoming). Mirror neurons and social cognition. *Mind and Language*.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children’s understanding of deception. *Cognition*, *13*(1), 103–128.