Dynamic Embodied Cognition

Leon C. de Bruin · Lena Kästner

Abstract In this article, we investigate the merits of an enactive view of cognition for the contemporary debate about social cognition. If enactivism is to be a genuine alternative to classic cognitivism, it should be able to bridge the "cognitive gap", i.e. provide us with a convincing account of those higher forms of cognition that have traditionally been the focus of its cognitivist opponents. We show that, when it comes to social cognition, current articulations of enactivism are—despite their celebrated successes in explaining some cases of social interaction—not yet up to the task. This is because they (1) do not pay sufficient attention to the role of offline processing or "decoupling", and (2) obscure the cognitive gap by overemphasizing the role of phenomenology. We argue that the main challenge for the enactive view will be to acknowledge the importance of both coupled (online) and decoupled (offline) processes for basic and advanced forms of (social) cognition. To meet this challenge, we articulate a dynamic embodied view of cognition. We illustrate the fruitfulness of this approach by recourse to recent findings on false belief understanding.

Keywords Social cognition · Enactivism · Embodied cognition · Cognitive gap · False belief understanding

Introduction

Since the 1950s, artificial intelligence (AI) has been the "intellectual heart" (Boden 2006) of cognitive science. With the computer metaphor of the mind, the view that minds are best understood as computational devices and cognition as system-internal information processing has gained popularity.[1] In the contemporary debate about

---

[1] For the lack of a better definition, we talk about cognition mainly as a kind of *information processing*. We do not require this process to involve symbolic representations per se, however. Also, we use "cognition" and "cognitive processing" interchangeably. We assume that whatever does cognitive processing is reasonably called a cognitive system or cognitive agent and that cognitive systems (or agents) are capable of cognition.

L. C. de Bruin (✉) · L. Kästner
Department of Philosophy II, Ruhr-University Bochum, Universitätsstr. 150, 44801 Bochum, Germany
e-mail: lcdebruin@gmail.com

how to best understand cognition, we can find two prominent proponents of this view: those committed to Good Old-Fashioned AI (GOFAI) argue that cognition is essentially rule-based computation over symbolic representations, while connectionist*s* claim it is computation over sub-symbolic or distributed representations that is not (or at least not explicitly) rule-based. Despite their differences, both views exemplify *classic cognitivism*.

> Classic Cognitivism (COG): The mind is basically an intracranial information processing system manipulating (sub-)symbolic representations; cognition essentially is this computational process.

Advocates of COG have traditionally maintained that cognition is primarily an "offline" exercise of "mental gymnastics" (Chemero 2009). Although this exercise may (but not necessarily does) involve a physical body and is often part of an elaborate interaction with the external world (think of calculating the change you expect to receive in a shop), these facts are best taken to be of *secondary* importance. COG proponents thus emphasize that offline processing involves internal representations, which are not bound to the current features of the agent's body or her environment and hence "decoupled".[2]

Classic Cognitivism has had serious ramifications for the study of social cognition. For several decades, research on how we understand others has been conducted under the heading of "mindreading". Mindreading, as Baron-Cohen (2001) puts it, is the ability to "infer the full range of mental states (beliefs, desires, intentions, imagination, emotions, etc.) that cause action" (p. 174) and use these inferences to predict and explain the behavior of others. There are two dominant approaches to explaining mindreading: theory theory (TT) and simulation theory (ST).

According to TT, mindreading is enabled by a folk psychological theory that specifies how mental states (in particular beliefs and desires) interrelate and give rise to intentions and actions. There are various versions of TT, for example "modular TT" (Fodor 1992; Leslie et al. 2005), "scientific TT" (Gopnik and Meltzoff 1997), "model TT" (Maibom 2003, Godfrey-Smith 2005) and "external TT" (Braddon-Mitchell and Jackson 2007). ST is usually portrayed as the main rival of TT because it denies that we need a folk psychological theory to understand others. Instead, ST claims mindreading involves putting ourselves in someone else's shoes by simulating their mental states while adjusting for the relevant differences. Like TT, ST has been developed into several directions; among them "explicit" ST (Heal 1986; Goldman 1989), "radical" ST (Gordon 1986, 2008), and "implicit" ST (Hurley 2008; Gallese 2005). Nowadays, most proponents of TT and ST favor hybrid models that accommodate both theorizing and simulation.[3] However, they still consider offline mindreading to be the central explanandum of social cognition.

In recent years, there has been a growing criticism of TT and ST. Several authors have argued against the importance of mindreading for social cognition and challenged the classic cognitivism underneath (e.g., Hutto 2004, 2008; Gallese

---

[2] In "Dynamic Embodied Cognition" section we provide a more detailed discussion of the terms "offline processing", "representation" and "decoupled".
[3] See Goldman (2006), Currie and Ravenscroft (2002) and Nichols and Stich (2003) for ST/TT hybrid accounts.

2005; Gallagher 2004, 2005, 2007; Ratcliffe 2006, 2007; Fuchs and De Jaegher 2009). Many of these critics have promoted the importance of the agent's body and the surrounding environment for cognitive processes under the heading of *enactivism*. Already proponents of embodied, embedded, and even extended cognition assign functional or implementational roles to the agent's extra-cranial body and her environment; however, enactivism goes beyond that: its proponents claim that cognition is *enacted* rather than residing somewhere in the agent's head or the world. In particular, they have claimed, cognition is not mere manipulation of (sub-)symbolic representations but an *interactive* process of sense-making—a *relational* activity between agent and environment. The term "enactivism" can be traced back to "The Embodied Mind", where Varela, Thompson and Rosch (1991) used it to designate a new way of thinking about the mind. Complaining that classic cognitivists left out "what it means to be human" (p. xv), they advocated focusing on the experiential nature of cognition as it arises from dynamic sensorimotor coupling and reciprocal determination between organism and environment. This characterization is still prominent:

> According to the enactive approach in cognitive science, cognition is grounded on the sense-making activity of autonomous agents—beings that actively generate and sustain themselves, and thereby enact or bring forth their own domains of meaning and value […]. (Thompson and Stapleton 2009, p. 23)

Although the enactive approach is sometimes celebrated as a "new paradigm" for cognitive science (e.g., Stewart et al. 2011), it is far from being unified. Rather, as we will show below ("Dynamic Embodied Cognition" section), enactivism turns out to be an umbrella under which a range of related approached find shelter. For now, however, it suffices to describe enactivism as follows:

> Enactive Cognition (ENAC): Rather than a representational process, cognition is a process of sense-making that emerges from the dynamic online interaction or 'coupling' between autonomous agents and the environment which they are embedded.

ENAC conceives of cognition as a process of sense-making emerging from the dynamic online interaction between agents and environment. As such, it can conveniently be modeled by *dynamical systems theory* (DST). According to DST, cognitive systems are best characterized as sets of differential equations the variables of which dynamically change their values over time, i.e. they evolve, in accordance with a set of dynamical laws (e.g., Port and Van Gelder 1995; Spivey 2007; Van Gelder 1995, 1998). Cognition thus is best understood as state–space evolution within a multidimensional dynamical system where the number of dimensions is given by the number of evolving variables.[4]

---

[4] DST provides a mathematical model using mathematical vocabulary, and can be used to model various processes in brain, body, and environment using a unified vocabulary. Since this makes DST attractive for proponents of extended cognition, it is often appealed to when cognition is thought to go beyond cranial boundaries. We will take advantage of this feature of DST when spelling out our own position ("Dynamic Embodied Cognition" section) as its unified vocabulary allows to capture immediate presentations, (sub-) symbolic representations, as well as intermediate kinds of stand-ins.

In this article, we investigate the merits of an enactive conception of cognition for the contemporary debate about social cognition. If ENAC is to be a genuine alternative to COG, it should be able to bridge what De Jaegher and Froese (2009) have called the "cognitive gap", i.e. provide us with a convincing account of those higher forms of cognition that have traditionally been the focus of its cognitivist opponents. In the next section, we show that, at least when it comes to social cognition, current articulations of enactivism are not yet up to the task. This is because they (a) do not pay sufficient attention to the role of offline processing or *decoupling* for social cognition, and (b) overemphasize phenomenology at the cost of embodiment. The main challenge for any enactive account of social cognition, so we argue, will be to acknowledge the importance of both coupled (online) and decoupled (offline) processes in advanced but also basic forms of social cognition. To meet this challenge, we develop a dynamic embodied view of (social) cognition ("Dynamic Embodied Cognition" section). We illustrate the fruitfulness of this view by applying it to recent findings on "implicit" false belief understanding ("The development of false belief understanding" section).

The cognitive gap

The enactive approach and the appeal to phenomenology

When the basic principles of the enactive approach are applied to social cognition, an agent's interaction with the world—including other agents—is transformed into what De Jaegher and Di Paolo (2007) call "participatory sense-making": "the coordination of intentional activity in interaction, whereby individual sense-making processes are affected and new domains of social sense-making can be generated that were not available to each individual on her own" (pp. 12–13; cf. Fuchs and De Jaegher 2009; De Jaegher et al. 2010). Thus, from an enactive point of view, the proper unit of analysis in social cognition is not the individual agent or the individual brain, but the coupled system as a whole; i.e. the participants, their dynamic interactions, and the context in which these interactions take place. What is characteristic of such a dynamic "coupled systems" view of social cognition is its emphasis on reciprocal interaction and recurrent feedback loops. It is precisely the *active* process of engaging with others that *constitutes* social cognition. What is appealing about the enactive view is the possibility to account for our everyday interactions without having to invoke the mindreading procedures so cherished by proponents of TT and ST. In other words, the enactive approach promises to explain social cognition in a direct and non-representational way, without folk psychological theories or simulation routines.

Proponents of ENAC frequently draw on phenomenological considerations in their criticism of COG (e.g., Varela et al. 1991; Gallagher and Zahavi 2008). Gallagher (2007), for example, has made the following objection against Goldman's claim that simulation is the "fundamental, default procedure" (2006, p. 175) of how we understand others. If simulation routines are indeed employed in a frequent and explicit fashion, then we should be aware of the different steps that we go through as we consciously simulate the other's mental states. However, according to Gallagher,

when I interact with others and try to understand them, "there is no experiential evidence that I use such conscious (imaginative, introspective) simulation routines" (2007, p.65).

In response, Goldman (2006) has argued that the appeal to phenomenology is problematic because phenomenology is "incapable of supporting weighty theses", hard to agree upon and "hotly disputed" (p. 249). Some have taken this criticism a step further. For instance, Spaulding (2010) assumes that neither the process of mindreading (theorizing or simulating) nor its product (the explanation or prediction of behavior) need be consciously accessible or phenomenologically transparent. Consequently, she claims that "the fallibility of phenomenology is one reason to doubt Gallagher's phenomenological argument. The total irrelevance of phenomenology is another" (p. 131).

For reasons of space we will not discuss the role of phenomenology in the debate on social cognition in detail. What seems to be important at this point is this: it is unclear how the appeal to phenomenology can yield conclusive answers to questions about the frequency or pervasiveness (i.e. it being the default mode) of mindreading—be it theorizing or simulation. This rather seems to be an empirical question.

Spauldings' criticism of enactivism

Despite the problems with phenomenological reflections, ENAC supporters typically use them as guidance in their approach to (social) cognition. Gallagher (2007), for instance, concludes that we usually do not need to consult a folk psychological theory or run a complicated simulation routine because phenomenology shows that in our everyday interactions we directly perceive and respond to meanings in the other's action. The immediate and responsive "online" nature of our social interactions is frequently explicated in terms of basic *embodied practices* (e.g., intentionality detection, shared attention, the perception of meaning and emotion in movement and posture). Proponents of ENAC argue that there are basically two ways in which such practices are primary to the "offline" modes of mindreading advocated by TT and ST. First, online social interaction involves abilities that are *developmentally* more fundamental. Call this the *argument from developmental primacy*. Second, online social interaction is primary to offline social interaction in the sense that it continues to characterize most of our participatory sense-making, and remains *the default or pervasive mode* of how we understand others (Gallagher 2001, 2005). This argument depends on the *appeal to phenomenology*.

Spaulding (2010) argues that both arguments are untenable to make her case against what she calls "Embodied Cognition" (EC)—her label for ENAC. Spaulding rejects the phenomenological argument against the pervasiveness of mindreading because it fails to establish that mindreading occurs only in rare circumstances (see the previous paragraph). Her criticism of the developmental primacy heavily relies on a distinction between two versions of this claim: according to the strong version, children are completely incapable of any social interaction before they master the basics of mindreading; according to a weaker version, children's social understanding is limited (but not inexistent) before they learn to robustly attribute mental states to others. Spaulding argues that the stronger version is obviously false, and it is hard

to disagree with her on this point given the social interaction even newborns can engage in. However, Spaulding thinks the weaker version (which is the view actually defended in the literature) is also false. This is because even proponents of COG can happily acknowledge the existence of certain low-level mechanisms ontogenetically primary to "proper" mindreading abilities—the "precursors" to a full-blown Theory of Mind (cf. Currie 2008). On this line of argument, the central difference between COG and ENAC seems to boil down to a different interpretation of low-level early mindreading abilities: while liberal proponents of COG see them merely as a transition stage (this is Spaulding's view), proponents of ENAC see them as the primary mode of social interaction even in adulthood.

While we agree with many of Spaulding's observations, we think that her conclusion is unsatisfactory because it leaves intact a fundamental dichotomy between COG and ENAC. This is not a problem *as such*, but it seems to result in a strictly separate treatment of mindreading and non-mindreading abilities (the "precursors") that fails to address the bigger picture and might not be adequate given the development of social cognition (see "The development of false belief understanding" section). In the remainder of this section, we will show that proponents of the enactive approach in fact face a similar problem, in the sense that they neglect to give a convincing explanation of their opponents' target explanandum.

The cognitive gap

Although the arguments from phenomenology are often used to criticize the pervasiveness of mindreading, its enactive wielders usually do not deny that we sometimes *do* engage in specialized TT or ST procedures. The point is that this happens only when our everyday agent–agent interactions fail or break down. Thus, they say, "such specialized cognitive approaches do not characterize our primary or everyday encounters with others" (Gallagher 2004, p. 202).

However, quarrels about what is and what is not the *primary* mode of everyday social interaction seem to distract us from what is *really* at stake in the debate. On the one hand, there obviously are cases of social cognition (e.g., gaze following, joint attention, imitation) that are carried out online and for which there is a straightforward enactivist interpretation—one that may be cognitively less demanding since it does not rely on building up costly internal representations and therefore is evolutionarily sensible.[5] On the other hand, there clearly are also cases of social cognition that are carried out offline and for which classical ST/TT accounts seem much more plausible (e.g., think of explicit forms of first-order and second-order false belief understanding). The problem is not simply that mindreaders deny the former and enactivists deny the latter. Rather, it seems both camps just (over-)emphasize the importance of offline and online social cognition, respectively.

This dichotomy does not further the debate on social cognition. Instead, it results in a divide and conquer strategy with mindreaders trying to account for advanced social abilities in terms of higher-order offline representational abilities, and enactivists focusing on more elementary forms of social interactions in terms of

---

[5] See, for example, Bermúdez (2003, pp. 31–2), and Spaulding (2010) for a discussion of this argument.

dynamic online couplings between agents. This is not only unsatisfactory for those in favor of mindreading, as we suggested above, but also for promoters of enactivism; for each neglects what the other focuses on.

ENAC has been celebrated as a revolutionary new paradigm replacing dusted COG. However, it cannot quite keep this promise and offer a genuine alternative to COG as long as it ignores the traditional explanandum of COG. That is, in order to replace COG, ENAC must not only account for dynamic online cognition but also for the "mindreading" abilities that have traditionally been the focus of TT and ST. This is so even if such abilities are only recruited in exceptional situations. The perhaps biggest—and yet unresolved—challenge for enactivism is to bridge this "cognitive gap", and provide us with a convincing account of offline social cognitive capacities. Thus, ENAC has to show how "an explanatory framework that accounts for basic biological processes can be systematically extended to incorporate the highest reaches of human cognition" (De Jaegher and Froese 2009, p. 439).[6] Problematically, though, phenomenological arguments tend to obscure this point. Hence, there are two serious challenges for ENAC: (1) restrict the emphasis on phenomenology to get a clear view of the cognitive gap, and (2) bridge the cognitive gap.

If ENAC is to bridge the cognitive gap, it must not only acknowledge advanced offline social cognitive abilities but also explain these "decoupled" modes of social cognition without falling back into the cognitivist camp. That is, the enactivist has to tell a story about how offline (i.e. decoupled) social cognition is grounded in and emerges from online (i.e. coupled) interaction.

To conclude our diagnosis: neither COG nor ENAC has been successful in providing a convincing account of both online and offline forms of cognitive processing. It hence seems fruitful to aim at a unified theoretical framework that solves the stalemate between ENAC and COG and integrates online and offline processes into a coherent story of how cognition can best be understood. This is our aim for the remainder of this article.

Dynamic Embodied Cognition

Two versions of enactivism

We already mentioned that enactivism is far from being a fully developed research program; it still lacks a unifying theoretical, conceptual, and methodological characterization. Without delving into too much detail, it will be important to at least distinguish between two readings of ENAC[7]:

---

[6] Note that the principal problem is not limited to social cognition but generalizes to other cognitive processes: "What, in general, is the relation between the strategies used to solve basic problems of perception and action and those used to solve more abstract or higher level problems?" (Clark 2001, p. 135).

[7] Although similar proposals have been made before (Walter 2010; Torrance 2006), not much attention is paid to this distinction in the literature. We believe that this may be one of the reasons why ENAC is so difficult to grasp.

Broad Conception of Enactivism ($ENAC_b$): Cognitive Science is concerned with the question "What is an agent with a mind?"; cognition is the relational process of *sense-making* emergent from the agent's dynamic coupling to her surroundings, an autonomous process that self-sustains the agent.

and

Narrow Conception of Enactivism ($ENAC_n$): Cognitive Science tries to characterize the nature of perception and perceptual experience; cognition is grounded in exploratory activities giving rise to perceptual consciousness.

$ENAC_b$ is a generic view that focuses on *autopoetic* (i.e. autonomously self-organizing) systems and how they are maintained. Its defenders include, among others, Thompson (2005, 2007) and Maturana and Varela (1980). Proponents of $ENAC_n$, on the other hand, put much more emphasis on how we become conscious of our immediate surroundings. They reject that our experiences result from manipulating detailed internal representations. Instead, they suggest a skill-based view according to which agents use a certain skill set to pick up details from the environment as needed—something that is familiar from embodied and embedded approaches to cognition (e.g., Clark 1989). What sets apart $ENAC_n$ from these views, however, is its expressed focus on conscious experiences, particularly perceptual consciousness. $ENAC_n$ is famously associated with O'Regan and Noë's work on visual perception (O'Regan and Noë 2001; Noë 2004) who argue that perception is essentially skillful explanatory activity—mere stimulation, i.e. sensory input, is insufficient for us to perceive and become aware of what lies in front of us; instead, practical knowledge (so-called *sensorimotor contingencies*) relating our action to changes in stimulation is required to make sense of the input.

Studies like O'Regan and Noë's are, however, not unequivocally perceived as evidence for cognition being enactive. On the one hand, they are also cited to support embodied and embedded views of cognition; on the other, proponents of $ENAC_b$ may deny research within the narrower framework to properly count as enactivism because it does not offer as radical an alternative to COG as does the broad conception. If we consider the historical development of both positions, it becomes clear why there is this divide within the enactivist camp: while $ENAC_b$ is commonly read as descending from *ecological psychology*, $ENAC_n$ may be conceived of as situatedness taken to the extreme—either as an intense version of the extended view (Robbins and Aydede 2009) or a combination of both extreme embeddedness and extreme embodiment (Rupert 2009 seems to assume something like this). In both cases, $ENAC_n$ descends from COG since situated views are probably best understood as attempts to liberate COG by adding environmental and bodily features to the classical representation-based internal processing. What sets apart $ENAC_n$ from COG, however, is its tendency to focus on active online rather than internal offline processing. Nevertheless, $ENAC_n$ bears a strong cognitivist heritage: it still conceives of cognition as a computational process—albeit one that does involve things other than internal (sub-)symbolic representations—and hence simply "waters down" (Chemero 2009, p. 30) COG. Read this way, $ENAC_n$ cannot bridge the cognitive gap but will be doomed to stay on the cognitivist side. Therefore, opting for a narrow conception of ENAC will not do the trick if one aims to bridge the cognitive gap.

ENAC$_b$ rejects COGs central tenets, i.e. the computational nature of cognition and the importance of some kind of representational structure. In return, it directs all its attention towards dynamic coupled interactions between agents and their environments which is accompanied by an almost complete neglect of offline processes. This in turn results in a failure to address the cognitive gap. Opting for a broad conception of ENAC thus will not enable us to bridge the cognitive gap either. As illustrated in "The cognitive gap" section, the emphasis of the experiential nature of cognitive processing in both versions of ENAC obscures these issues. In order to develop a view that can bridge the cognitive gap, we need to not only restrict the emphasis on phenomenology but revise ENAC$_b$'s and ENAC$_n$'s conception of an agent's dynamic coupling to her surroundings and bring (at least some) capacities for offline processing back into the picture. This will also require us to reconsider the role of representations in cognitive processing.

Dynamic Embodied Cognition

This section is devoted to developing a view that explicitly takes the problems with ENAC$_b$ and ENAC$_n$ on board: *Dynamic Embodied Cognition* (DEC). DEC can be seen as an attempt to offer a framework that spans the whole spectrum from COG to ENAC and offers a way of bridging the cognitive gap. Since our starting point is an embodied view of cognition, DEC may also be interpreted as a dynamicist version of embodiment.

What sets apart DEC from enactivism is the recognition that an agent can—to a certain degree—*decouple* from her environment and is not entirely dependent upon direct online interaction. On the one hand, as we pointed out above, ENAC$_n$ proponents tend to hold that *all* cognition is grounded in exploratory interaction with the environment. Although DEC does not deny an essential role for these activities, and in so far is in tune with ENAC$_n$, it does not take these exploratory activities to be *all* there is to cognition—DEC explicitly allows for decoupled processes to play a substantial role as well—and hence takes a crucial step away from ENAC$_n$. Supporters of ENAC$_b$, on the other hand, tend to see dynamic coupling between agent and environment as ultimately replacing the representational structures postulated by their cognitivist opponents. In fact, a lot of effort has been put into getting rid of the notion of "representation" altogether (e.g., Gallagher 2008; Garzón 2008; Hutto 2008, 2011). However, an integrative theoretical framework will have to deal with cases where cognitive processing, at least to some extent, takes place independently of the agent's environment. Sometimes an agent simply cannot rely on what is right there, and she will have to appeal to some kind of stand-in for what is currently absent. This may be particularly relevant in social contexts where agents do not always have direct access to the mental states (e.g., beliefs, desires) of others. Furthermore, an agent may also purposefully distance herself from her immediate surroundings to not automatically act upon particular affordances. She may instead consider other ways to respond to the affordances the environment offers, contemplate upon various possibilities for actions towards a certain goal, etc. That is, the agent may engage in "offline" processing, while *controlling for* or *suppressing* "online" processing, to achieve the best result. Focusing on direct coupling (online processing) alone neglects the agent's ability to generate *new* and

more *advantageous* conditions for relating to her environment by engaging in offline processing.[8]

What is characteristic of offline processing is the manipulation of information that is currently unavailable from the environment (but typically has been present on previous occasions), and therefore has to be internally represented. We contrast this type of processing with online processing, which does not require recourse to such internal representations but instead manipulates what is currently present in the environment and readily available to the agent. The distinction we draw between the information used for online and offline processing parallels Shannon's (1993) distinction between "presentations" and "representations". Presentations are what is currently within the agent's grasp and can be readily used in forms of processing in which the agent is directly coupled to her environment. Representations, on the other hand, are not directly bound to what lies in front of the agent (although they otherwise bear characteristics similar to those of presentations); they are decoupled from the environment—although typically *arising from* or *grounded in* enaction before taken offline—and thus suitable for offline processing.[9] The agent may engage in both types of processing differentially depending on the environment she is situated in, and the internal processing resources at her disposal. When the agent is able to represent the information missing from the environment (e.g., based on memory), she may be able to achieve what under other circumstances she could accomplish by online processing alone—or even engage in more sophisticated tasks unsuited for online processing.

Decoupling, as we see it, is a matter of reducing direct effects of environmental stimulation and opening up possibilities for internally regulated behavior. This enables the agent to inhibit reflex-like automatic responses to external triggers (presentations), and elicit actions in the absence of an external trigger, for example by recruiting internal representations in scenarios that do not provide the presentations triggering coupled online processing. To illustrate this, think of the walking reflex: as soon as their soles touch a flat surface, newborns attempt to walk by placing one foot in front of the other. The infant's behavior is purely automatic; it can neither be inhibited, i.e., the environmental input cannot be ignored and the newborn will engage in walking movements even lying on its back with its feet pointing upward, nor can walking behavior be triggered without a flat surface under the infant's soles. This seems to be a perfectly fine example of direct dynamic coupling between agent and environment. The infant does not have the capacity to walk independently of what touches the sole of her feet; her behavior thus is not *decoupled*, it is not independent of environmental stimulation. At about 8 weeks of age, though, the infant acquires the ability to inhibit the walking reflex—she learns to decouple her behavior and the walking reflex disappears as an automatic response.

---

[8] This is not to say, however, that any kind of offline processing requires the kind of (sub-)symbolic representationalism advocated by COG. Rather, as we argue in the "Dynamical system theory" section, when talking about internal representations being recruited for offline processing we think of any kind of internal stand-in for something currently not present in the environment. Such stand-ins, we suggest, are best conceptualized as primarily *grounded in* enaction but then taken offline.

[9] This is not to say representations can only be formed through direct contact with their referent, however. An agent may acquire representations of abstract entities and things she never got in touch with, e.g., by use of imagination or drawing analogies to things she did encounter (cf. Lakoff and Johnston 1980).

Towards the end of the first year, the infant acquires the ability to walk, that is she becomes able to initiate walking as voluntary behavior—walking is now fully decoupled, and has become (at least to some extent) independent of what kind of surface touches her soles (see, e.g., Adolph et al. 2003 for a more detailed discussion).

This is a good example of how decoupling provides an agent with a certain degree of *autonomy* towards the presence of bodily and environmental features. The child acquires the capacity to (a) inhibit automatic behavioral responses (walking motion) vis-à-vis environmental features (a flat surface touching their soles) and (b) elicit that behavior (formerly coupled to the presence of a certain stimulus) independently of environmental stimulation. Analogous observations can be made in the development of social cognition. For example, several researchers report that young infants are like "imitation machines" in the sense that they copy and reproduce virtually all human actions they see in their surroundings (Carpenter et al. 1998; Tomasello 1999). Although these infants cannot help but respond to what they perceive, they do have the capacity to correct and improve their imitations. Moreover, with age they come to master forms of imitation which are beyond the scope of younger infants. By 18 months, for instance, infants can complete unfinished actions of agents they observe, such as pulling apart miniature dumbbells (Meltzoff 1995; Meltzoff and Brooks 2001). In this respect, imitation in infants quickly develops into more than a mechanical reflex-like response—it starts to involve a kind of responsiveness that is best characterized as a dynamic process of coupling and decoupling. Although imitative behavior is triggered by the perception of other agents it is not fully *determined* by this action–perception coupling but (at least partly) independent, thus *decoupled* or *autonomous*, of what is going on around the agent.

In the enactivist literature, the emphasis on autonomy is prominent. An autonomous agent is sometimes defined as an autopoietic system that (a) generates its own identity and the particular conditions by which it can relate to its environment (the system has an intrinsic teleology and constitutes its own purposeful and goal-directed existence), and (b) makes sense of the affordances offered by the environment in relation to its particular way of realizing and preserving its identity (the system has a capacity for self-monitoring and appropriate regulation, and can adapt itself to changes in the environment). This definition is demanding and it is not unequivocally shared among enactivists (see Weber and Varela 2002; Froese and Di Paolo 2009, and Barandiaran et al. 2009 for further discussions of autonomy). For current purposes, since we are not concerned with issues of self-organization and self-sustainment, it suffices to use a simpler, less demanding understanding of autonomy that still captures the effects of decoupling. "Autonomy" as we will use it simply refers to a certain *degree of independence* of environmental stimulation and the ability to inhibit and elicit actions that are typically directly coupled to the perception of certain environmental features and/or the affordances they offer.[10]

The fact that our weaker definition of autonomy allows for degrees of independence is crucial, for there certainly seem to be different extents to which

---

[10] The terms 'decoupling' and 'autonomy' are used in a similar sense by Barandiaran and Moreno (2008) in modeling adaptive autonomous agents.

systems are autonomous or decoupled from their environments.[11] For instance, an agent may be able to inhibit automatic responses without being able to initiate certain actions in the absence of environmental input. This makes her more autonomous than an agent who is unable to inhibit automatic responses, but less autonomous than an agent who can additionally elicit actions in the absence of environmental stimulation. Importantly, degrees of autonomy and decoupling are always relative to a given environment. The more an agent's action becomes independent of the specific environmental features, the more she decouples from the environment and the more autonomous she becomes. However, a significant change in the environment can have a dramatic impact on the agent's autonomy and her abilities for offline processing. For example, healthy and mature human beings will generally have little difficulty to initiate walking on a firm surface. However, put them in a radically different environment, e.g., on the bottom of the ocean or the surface of the moon, and things start to change rapidly; at least unless means for dealing with these new environments—such as special equipment—are provided.

The coupling and decoupling relations between agent and environment advocated by DEC are *dynamic* in the sense that they are a matter of degree and never an end in itself. On the contrary, decoupling mainly services *recoupling*: taking certain processes offline, decoupling allows for novel ways of relating to the environment that provide the agent with more and better tools to act on certain affordances or create new ones. This, in turn, enables the agent to re-enter into online processing. The dynamic interplay of decoupled and coupled processes may be used for optimization of cognitive processing. The obvious (evolutionary) advantage of this cognitive flexibility and autonomy from environmental stimulation is that agents become less dependent upon, and gain new ways of relating to, their direct surroundings, including (in the case of social cognition) their conspecifics. Thus, the development from online to offline forms of cognitive processing is to be seen as one of progressive independence from environmental and bodily features, giving the agent autonomy but at the same time demanding more cognitive work. Against the background of developmental studies, this is a plausible story of how the developing organism progressively acquires decoupled abilities and thus engages in offline cognition while online cognition will still be used throughout adulthood where feasible. We will come to illustrate these points in our discussion of empirical studies on social cognition ("The development of false belief understanding" section).

If the recourse to offline processing advocated by DEC is useful the strict non-representationalism built into $ENAC_b$ is too strong. At the same time, however, COG's strict representationalism also seems too demanding for the kind of offline processing we have in mind.[12] Therefore, we propose a weaker reading of

---

[11] De Jaegher and Di Paolo (2007) seem to share this intuition: they propose that the autonomy of agents involved in participatory sense-making can be "augmented or reduced" (p. 493) under particular circumstances.

[12] COG proponents tend to have a rather strong reading of "representation", according to which representations are either *symbolic* (in the case of GOFAI) stand-ins that map, in a one-to-one fashion, onto external objects or (in the case of Connectionism) sub-symbolic structures mapping onto external items. In a nutshell, sub-symbolic structures are composed of multiple components whereof any one component can, but does not have to, figure in multiple other representations, which in turn may include additional components.

"representation" according to which a representation can be *any* kind of stand-in for another item that will typically be best understood as *grounded in* enaction but subsequently taken offline. In this sense, there does not have to be a direct mapping between internal stand-ins and external items; the reading from the barometer represents the current air pressure just as a certain neural activation pattern in the brain represents face perception and the word "cat" represents the concept of a cat.

This more liberal conception of representations (a) still picks out something very different from Shannon's presentations, and (b) does not *ipso facto* exclude (sub-) symbolic representations but instead widens the scope of what counts as a representation. As such, it allows us to take a middle course between COG, $ENAC_n$, and $ENAC_b$. We criticized $ENAC_b$'s attempt to discard representational structures altogether in the light of the cognitive gap. If we aim to bridge that gap, we take it, some form of representational processing will be required. In order to achieve this, DEC allows for decoupling from the agent's environment. Our understanding of "autonomy" as independence of environmental features is, although inspired by $ENAC_b$, different from the standard enactivist reading. This marks an attachment point between DEC and $ENAC_b$ although at the same time it clearly sets them apart. We already pointed out that, due to its cognitivist heritage, $ENAC_n$ is in principle compatible with the computational nature of cognitive processing. DEC reconciles COG and $ENAC_n$ by leaving the assumption of cognition being some form of computational process untouched and allowing immediate presentations and (sub-) symbolic representations of environmental features as well as other kinds of stand-ins to be recruited for this process.[13] Rather than pure enactment of the world, we therefore advocate a mixture of enactive and representational processing that may best be captured within DST.[14]

Drawing together the threads of the current section, we can define Dynamic Embodied Cognition as follows:

> Dynamic Embodied Cognition (DEC): Cognition is (developmentally) grounded in the agent's coupling to her surroundings. This coupling is dynamic insofar it allows agents to (a) rely on direct online cognitive processes, or (b) decouple from their environment and engage in offline cognitive processing. Online processing is 'cheap' and efficient as it allows the agent to avoid building up 'costly' internal representations which would require additional processing resources; however, it comes at the cost of limited and inflexible responsiveness. Offline processing provides agents with more flexibility (autonomy) regarding their direct environment, but is also cognitively more demanding. For a system to engage in decoupled offline

---

[13] This does not mean that DEC should simply be seen as an extended version of $ENAC_n$. The primary difference is that DEC emphasizes the importance of decoupled cognition, something that is largely ignored (if not denied) by $ENAC_n$. Nevertheless, it is probably true that DEC comes closer to reconciling $ENAC_n$ and COG than $ENAC_b$ and COG.

[14] The fact that ENAC and DEC agree that DST provides a convenient model of cognitive processing does *not* entail the two views collapse; for, as we see it, DST is not an account of cognition in its own right but a mathematical model of cognitive systems and cognitive processing. There is an important difference between the dynamic models of cognition on both views: while ENAC appeals to DST *only* to build highly dynamic models, DEC appeals to DST as a framework able to model both dynamic and more static processes with fewer variables.

processing, it has to (a) inhibit automatic behavioral responses to environmental features and (b) be able to elicit the behavior formerly coupled to the presence of a certain stimulus independently of environmental stimulation.

Having removed the obscuring emphasis on phenomenology characteristic of ENAC, we are now in a position to recognize that bridging the cognitive gap will require us to acknowledge both online and offline processes within a unified theoretical framework. In order to achieve this, DEC allows for a dynamic interplay of coupled and decoupled processes. This, in turn, requires a more liberal notion of representation than advocated by COG. However, (sub-)symbolic representational processing is still an option within DEC, just a very specific one that may have only limited application—perhaps in later development or higher-order cognition. This way, the traditional explananda of COG do not fall out of the spectrum and the cognitive gap can be bridged.

Dynamical Systems Theory

If DEC is right and cognition really is best understood as the dynamic interplay of enaction and decoupled representational processing, what kind of architecture could be used to capture these facts and model cognitive systems? We here propose that Dynamical Systems Theory may be our best bet.

DST models cognitive systems as sets of differential equations the variables of which dynamically change their values over time. Over time, the dynamical system evolves through a multidimensional state space. The resulting trajectories through state space may be said to represent (in a weak sense) the system's state.[15] "Computation" in a dynamical system no longer refers to discrete steps of symbol manipulation but should be conceived of as transforming one kind of information into another where this transformation may occur by dynamic evolution through a system's states.

We agree with Spivey ([2007](#)) that the "traditional information-processing approach (borrowed from the early days of computing theory), […] place[s] too much emphasis on easily labeled static representations that are claimed to be computed at intermittently stable periods of time." (p. 4) Yet, we also think that—at least for offline processes—there has to be some kind of "mediating stand-in […] in between sensory stimulation and physical action" (ibid., p. 2). Our continued use of the term "representation" for these stand-ins serves to ease the "intellectual transition" (ibid.) from the classic information-processing framework of COG to a DST framework to model cognitive processing. The clue with using DST as a modeling framework is that it provides the tools to not only capture highly dynamic features of a system but also (although this is not the typical area of application for dynamical systems) static ones by introducing more static variables hence introducing stability and independence of, e.g., environmental factors.

Within DST, the decoupling we described above as a transition from online to offline processes will be modeled by reducing the number of dimensions of the dynamical system in question. That is, in the decoupled system, fewer variables will

---

[15] Note that although trajectories through state space may be said to represent the state of the system as a whole, this does not entail anything about the kinds of representations (or presentations) featuring in the cognitive processes being modeled by a dynamical system.

dynamically change their values over time while more variables get assigned fixed values making the system more static. This can be achieved by, e.g., identifying linear relations between two variables such that one of them can be substituted with the product of a scaling factor and the other variable. Similarly, one may assign approximate values for certain variables if, say, their values change within a very restricted interval. Such mathematical procedures result in a reduction of dimensions of the dynamical system.[16] Taking this line of reasoning a bit further, appropriately constrained dynamical systems can be used as models of classical cognitivist systems (see Garzón 2008 for a similar proposal). Likewise, recoupling can be conceptualized as adding degrees of freedom to an existing dynamical system thus making it more flexible and dynamic. Put slightly differently, the thought is this: dynamical systems have high dimensionality, where each dimension corresponds to a variable the value of which evolves through time. Reducing the number of unknowns (i.e. variables with changing values) hence reduces the number of dimensions of space through which a dynamical system evolves; the lower the dimensionality, the less dynamical (and more rigid) the system.

Hence, applied in the right way, DST shows how there can be a unifying understanding of seemingly very different (i.e. online and offline) processes; and such an understanding is what DEC offers. Although the story DEC tells about decoupling is not the mathematical process of constraining a dynamical system, mathematical constraining may be an adequate *model* of what happens when a cognitive system decouples from environmental features. Since both static and dynamic aspects of a system—important for modeling decoupled offline and coupled online processes, respectively—can be captured using the same mathematical vocabulary within DST, it comes handy to model cognitive systems as characterized in an integrative framework like DEC.

As we hope has become clear throughout this section, DEC respects both enactive online processing accounts and classicist offline processing accounts of cognition. While it agrees with enactivism that cognition is best modeled using DST, DEC shifts the focus from experiential features prevalent in both $ENAC_n$ and $ENAC_b$ towards a focus on dynamicism. At the same time, DEC significantly imports from embodied and embedded views of cognition. These bear a substantial cognitivist heritage insofar as they conceive of cognition as a computational process; yet, they also emphasize the importance of bodily and environmental features for at least some cases of (online) cognitive processing. While advocates of $ENAC_n$ and $ENAC_b$ typically deny the importance of offline processing completely, DEC acknowledges the whole spectrum from pure online to pure offline processing and thus offers a way to bridge the cognitive gap. More than that, it assigns an essential role to decoupling and recoupling in both cognitive processing and development. Since this makes for a stark contrast to enactivism, we locate ourselves in the embodied cognition rather than enactivist camp. In the following section, we will use recent developmental findings on false belief understanding in order to illustrate how DEC allows us to resolve the disagreement between COG and ENAC concerning the superiority of offline versus online forms of social cognition.

---

[16] Izhikevich (2007, ch. 5.2) shows how to do this using the example of Hodgkin and Huxley's (1952) multidimensional model of the action potential and reduce it to a simple two-dimensional model.

The development of false belief understanding

Recent findings on false belief understanding

Proponents of TT and ST frequently appeal to developmental psychology to support their views on social cognition. In particular the "elicited response" false belief test (ER-FBT) in which infants are asked to give an explicit prediction of another agent's behavior on the basis of her false belief, has been a very popular choice. In the unexpected location ER-FBT, for example, children observe a protagonist who sees an object being placed in a certain location (Wimmer and Perner 1983; Baron-Cohen et al. 1985). The protagonist leaves, and the object is moved. When the protagonist returns, she mistakenly believes the object is still in its initial location. At this point, the children are asked to (verbally) predict where the protagonist will look for the object. Test results show that 3-year-olds typically give a wrong answer to this question, while 4-year-olds answer correctly. Many researchers have therefore concluded that false belief understanding does not emerge until 4 years of age (Flavell 2004; Sodian 2005; see Wellman 2002 for a review, and Wellman et al. 2001 for a meta-analysis). The controversy between TT and ST has mainly centered on the question whether these findings are indicative of the four-year-olds' mastery of a folk psychology (in the fashion of TT) or rather a set of simulation routines (appealed to by ST).

Meanwhile, advocates of enactivism have argued that the results of ER-FBTs do neither support TT nor ST, because the test is not representative of the full scope of social cognitive abilities. Gallagher (2004), for example, claims that ER-FBTs are designed to capture a set of very specialized cognitive abilities, which "put us in an observational mode and do not capture the fuller picture of how we understand other people" (p. 204). This line of reasoning has also been taken by developmental psychologists more widely. For instance, Bloom and German (2000) argue that the ER-FBT is an "ingenious, but very difficult task that taps (only) one aspect of people's understanding of the minds of others" (p. 30).

The ER-FBT indeed places strong demands on children's cognitive abilities (Bloom and German 2000; Carlson and Moses 2001). Recent investigations of false belief understanding have attempted to reduce these demands in order to see whether children might be capable of false belief understanding at an earlier age. Employing violation of expectation, anticipatory looking, and active helping paradigms, these "spontaneous response" false belief tests (SR-FBTs) no longer require children to explicitly state the protagonist's belief. Instead, their false belief understanding is inferred from the behavior they spontaneously produce (Baillargeon et al. 2010). On the basis of these studies it has been claimed that false belief understanding emerges much earlier; it has been reported in 25-month-olds (Southgate et al. 2007), 15-month-olds (Onishi and Baillargeon 2005), 13-month-olds (Surian et al. 2007), and even 7-month-olds (Kovács et al. 2010). This early manifestation of false belief understanding has been called "implicit" (Ruffman et al. 2001), because the infants are not explicitly aware of the false belief of the protagonist. Despite the fact that their behavioral responses indicate sensitivity to the protagonist's false belief, however, they answer incorrectly when asked for a direct statement of it (Clements and Perner 1994).

Whether or not the SR-FBT findings should indeed be interpreted in terms of false belief has been hotly debated (Perner and Ruffman 2005; Ruffman and Perner 2005; Csibra and Southgate 2006; Sirois and Jackson 2007; Herschbach 2008). On the one hand, ENAC proponents argue that we can explain the SR-FBT results without assuming genuine false belief understanding on the part of the infant (e.g., Hutto 2011). However, such interpretations tend to ignore the role of offline processing and do not address how explicit false belief understanding enters into the picture in later developmental stages. On the other hand, COG proponents claim that SR-FBT results *are* proof of genuine false belief understanding (Surian et al. 2007; Baillargeon et al. 2010). This, however, seems to give rise to a developmental paradox: if young infants already understand false belief, then why do they fail the ER-FBT?

By applying DEC to false belief understanding, we think that some of these problems can be avoided. In what follows, we show that rephrasing the issue in terms of offline and online processes provides us with an explanation of false belief understanding as an "implicit" ability that starts out as grounded in basic online processes, albeit already partly decoupled, and develops into a more sophisticated "explicit" ability that relies on offline processes to a much larger extent.

False belief understanding and decoupling

Most accounts of false belief understanding assume that infants have a default tendency to attribute their own (true) beliefs to other agents (Leslie et al. 2004) or to respond on the basis of their own knowledge (Birch and Bloom 2007; Carlson and Moses 2001; Mitchell 1996; Russell 1996). In order to pass the ER-FBT, so it is thought, infants have to be capable of decoupling, i.e. take their own reality-congruent perspective offline (Scott and Baillargeon 2009).

Baillargeon et al. (2010), for instance, argue that this decoupling ability is precisely what makes the difference between the task demands of ER- versus SR-FBTs: while the SR-FBT only requires false belief representation, the ER-FBT additionally involves response inhibition (when selecting a response, children must inhibit any prepotent tendency to answer the test question based on their own knowledge), and response selection (when asked the test question, children must access their representation of the other's false belief to select a response).

However, this seems to ignore the fact that the SR-FBT involves offline processing as well, albeit of a less demanding kind. Several SR-FBTs (Onishi and Baillargeon 2005; Surian et al. 2007; Song and Baillargeon 2008; Träuble et al. 2010; Kovács et al. 2010) show that very young infants already understand that the visual perspective of another agent can be different from their own. For example, Southgate et al. (2007) employed an anticipatory looking SR-FBT in which 25-month-olds observed how a protagonist witnessed a puppet bear who hid a ball in one of two boxes. Then the protagonist got distracted and turned away from the scene. Meanwhile, the bear removed the ball from its original hiding place. Southgate et al. (2007) found that most 25-month-olds correctly anticipated the protagonist's behavior (i.e. where she would search for the ball on her return) and looked at the location where she falsely believed the ball to be hidden.

Although the experiment by Southgate et al. (2007) does not require infants to deal explicitly with differences in belief, it does require them to process differences

between the visual information available to themselves and the visual information available to the other agent. This can only be accomplished offline, since the other's visual information is not directly available to the infant and needs to be represented by her. Therefore, already this SR-FBT involves a capacity for decoupling from one's own online processing of visual information and processing offline a representation of the visual information accessible to another agent. Yet, the roles of decoupling and offline processing are still limited. The infant largely relies on online visual information, and only has to process offline the other agent's representation of the location of a *single* object. More difficult versions of the SR-FBT place stronger demands on offline processing. For example, Song and Baillargeon (2008) conducted an experiment in which infants had to represent the visual representation of another agent with respect to *both* the location and the identity of *two* objects. Among SR-FBTs, we can thus distinguish between more or less demanding versions requiring more or less decoupling respectively.

Crucially, the ER-FBT takes this decoupling further. First, it requires the infant to deal with more abstract task elements (e.g., a cartoon or a short story instead of the performing real-life agents and objects featuring in the SR-FBT) that have to be processed independently of the infant's direct environment (like a desk). This makes the ER-FBT more difficult than its more "pragmatically natural" counterpart, as Bloom and German (2000) have pointed out.

Second, the ER-FBT requires infants to not only represent but *meta-represent*. That is, infants have to (a) process offline the other agent's propositional attitude towards the object (i.e., her false belief about its location), and (b) come up with a verbal prediction of the agent's behavior on the basis of her belief. The decoupling required in (a) has been demonstrated to place increasing demands on *executive functioning*: several studies have found robust correlations between ER-FBT performance and response inhibition (Perner and Lang 1999; Cole and Mitchell 2000; Carlson and Moses 2001), and ER-FBT performance and working memory (Carlson et al. 2002; Hala et al. 2003; Perner et al. 2002). Concerning (b), note that *verbal interaction* between infant and experimenter has been reported to contribute to infants' difficulties with the ER-FBT (cf. Southgate et al. 2007). Indeed, many studies have found strong correlations between linguistic competence and ER-FBT performance (Dunn et al. 1991; Astington and Jenkins 1999; Gale et al. 1996; De Villiers and De Villiers 2000; Watson et al. 2002; Farrar and Maag 2002). A number of hypotheses have been advanced to explain why children have more difficulty with FBTs that involve linguistic interaction. Some researchers propose that children need to master its semantics (Moore et al. 1990), whereas others argue that what is required is getting a handle on its syntactic structure (Hale and Tager-Flusberg 2003; Lohmann and Tomasello 2003). DEC is not committed to any of these hypotheses in particular. However, note that it is very well possible that the ER-FBT requires a stronger form of decoupling precisely *because* it involves language. There is much more to be said about how executive function, memory, linguistic and other abilities scaffold the development of false belief understanding. This, however, will have to wait for another occasion.

Given the SR- and ER-FBT findings, some have argued that we should consider the existence of two systems for false belief understanding. Apperly and Butterfill (2009), for example, postulate a cognitively efficient but inflexible "minimal"

Theory of Mind to explain the SR-FBT results, and a flexible but cognitively demanding "full-blown" Theory of Mind to explain the ER-FBT findings. Such a proposal fits well with DEC's emphasis on the development of offline processing as cognitively demanding, but also providing a greater flexibility. However, the exact relation between the two systems suggested by Apperly & Butterfill remains unclear—rather than a unified account, they suggest to keep whatever explanation we already have by making room for each in a separate system. Applying DEC may be enlightening in this context: it suggests online mindreading abilities to not simply be a precursor of (and supported by a different system than) a full-blown Theory of Mind but to be continuously used where possible. Whether or not this is correct is, of course, an empirical question we cannot settle here (but see Aschersleben et al. 2008, Sodian 2010 for discussions about the relationship between the abilities involved in SR- and ER-FBTs).

We hope to have shown in this section that (a) even very basic social capacities such as implicit false belief understanding already involve decoupling, (b) once DEC is adopted, there is no cognitive gap between implicit and explicit false belief understanding since both can be understood in terms of different degrees of decoupling, (c) the ER-FBT involves stronger offline processing than the SR-FBT. By acknowledging the importance of offline processing, DEC is able—unlike ENAC—to give an account of false belief development "all the way up". At the same time, DEC's focus on the interplay between online and offline processing allows it to avoid a number of potential problems for standard COG "top-down" analyses of this phenomenon in terms of strong belief representations.

One may object, of course, that we do not actually provide a full dynamical story, a concrete dynamical system, for false belief understanding here. Neither have we shown what exactly the variables in the coupled system are that get constrained once the system decouples. This, however, is simply not possible as long as we do not have a specific dynamical model to look at and know what its variables stand for. It has not been our intention to provide such a dynamical model here. Rather, our aim has been to draw attention to the potential that DST bears as a modeling tool, and to argue that it squares well with the philosophical view of cognition that we have developed in the previous sections.

As a bit of consolation, though, let us point to a dynamical model for the A-not-B task proposed by Smith and Thelen (2003). In this task, infants are required to track the location of a toy the experimenter puts in either of two locations several times in a row. On the critical trial, the experimenter then chooses the other location. Infants up to about 12 months of age continue to look for the toy in the first location on that trial, despite having observed the experimenter putting the toy in the second location. Although not a false belief task, the A-not-B task may also be interpreted as testing for a basic form of decoupling when infants are required to inhibit the reaching behavior trained in previous trials to conform to the new situation and successfully retrieve the toy from the new location. If the transition from a stage where infants fail to do so to a stage where they successfully decouple can adequately be modeled by DST then it seems plausible to assume this also works for the development of false belief understanding. Future research might be able to provide a concrete dynamical model for (the development of) false belief understanding. Although we are in no position to do that here, we may—drawing

on our above considerations of dimensionality of more or less constrained dynamical systems—speculate that coupled, fluid and flexible systems primarily supporting online cognition will best be modeled by high-dimensional dynamical systems with many variables. Whereas systems with some decoupling (those passing the SR-FBT) may best be modeled by assigning some of these variables fixed values, highly decoupled systems (those passing the ER-FBT) may best be modeled by assigning many or most variables fixed values resulting in heavily reduced dimensionality. Further research will be required to examine the extent to which the theoretical framework we proposed here can be used to account for the full range of socio-cognitive abilities.

References

Adolph, K. E., Vereijken, B., & Shrout, P. E. (2003). What changes in infant walking and why. *Child Development, 74*, 474–497.

Apperly, I., & Butterfill, S. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*, 953–970.

Aschersleben, G., Hofer, T., & Jovanovic, B. (2008). The link between infant attention to goal-directed action and later theory of mind abilities. *Developmental Science, 11*, 862–868.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relationship between language and theory-of-mind development. *Developmental Psychology, 35*, 1311–1320.

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

Barandiaran, X., & Moreno, A. (2008). Adaptivity: from metabolism to behavior. *Adaptive Behavior, 16*, 325–344.

Baron-Cohen, S. (2001). Theory of mind in normal development and autism. *Prisme, 34*, 174–183.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*, 37–46.

Barandiaran, X., Di Paolo, E. A., & Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior, 17*(5), 367–386.

Bermúdez, J. L. (2003). The domain of folk psychology. In A. O'Hear (Ed.), *Minds and Persons* (pp. 25–48). Cambridge: Cambridge University Press.

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*(5), 382–386.

Bloom, P., & German, T. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*, B25–B31.

Boden, M. (2006). *Mind as machine*. Oxford: Oxford University Press.

Braddon-Mitchell, D., & Jackson, F. (2007). *Philosophy of mind and cognition*. Oxford: Blackwell.

Carlson, S., & Moses, L. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032–1053.

Carlson, S., Moses, L., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11*, 73–92.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior & Development, 21*, 315–330.

Chemero, A. (2009). *Radical embodied cognition*. Cambridge: MIT Press.

Clark, A. (1989). *Microcognition: Philosophy, cognitive science, and parallel distributed processing*. Cambridge: MIT Press.

Clark, A. (2001). Reasons, robots, and the extended mind. *Mind and Language, 16*, 121–133.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development, 9*, 377–397.

Cole, K., & Mitchell, P. (2000). Siblings in the development of executive control and a theory-of-mind. *British Journal of Developmental Psychology, 18*, 279–295.

Csibra, G., & Southgate, V. (2006). Evidence for infants' understanding of false beliefs should not be dismissed. Response to Ruffman and Perner. *Trends in Cognitive Science, 10*, 4–5.

Currie, G. (2008). Some ways to understand people. *Philosophical Explorations 11*(3), 211–218.

Currie, G., & Ravenscroft, I. (2002). *Recreative minds: Imagination in philosophy and psychology*. New York: Oxford University Press.

De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences, 6*, 485–507.

De Jaegher, H., & Froese, T. (2009). On the role of social interaction in individual agency. *Adaptive Behavior, 17*, 444–460.

De Jaegher, H., Di Paolo, E. A., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences, 14*, 441–447.

De Villiers, J., & De Villiers, P. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 191–228). Hove: Psychology Press.

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Younblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecendents. *Child Development, 62*, 1352–1366.

Farrar, M. J., & Maag, L. (2002). Early language development and the emergence of a theory of mind. *First language, 22*, 197–213.

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly Journal of Developmental Psychology, 50*, 274–290.

Fodor, J. (1992). A theory of the child's theory of mind. *Cognition, 44*, 283–296.

Froese, T., & Di Paolo, E. A. (2009). Sociality and the life-mind continuity thesis. *Phenomenology and the Cognitive Sciences, 8*, 439–463.

Fuchs, T., & De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences, 8*, 465–486.

Gale, E., deVilliers, P., deVilliers, J., & Pyers, J. (1996). Language and theory of mind in oral deaf children. Paper presented at the *Boston University Conference on Language Development*, Boston, MA.

Gallagher, S. (2001). The practice of mind: Theory, simulation, or interaction? *Journal of Consciousness Studies, 8*, 83–107.

Gallagher, S. (2004). Understanding interpersonal problems in autism: Interaction theory as an alternative to theory of mind. *Philosophy, Psychiatry, and Psychology, 11*, 199–217.

Gallagher, S. (2005). *How the body shapes the mind*. New York: Oxford University Press.

Gallagher, S. (2007). Simulation trouble. *Social Neuroscience, 2*, 353–365.

Gallagher, S. (2008). Are minimal representations still representations? *International Journal of Philosophical Studies, 16*, 351–369.

Gallagher, S., & Zahavi, D. (2008). *The phenomenological mind: An introduction to philosophy of mind and cognitive science*. London: Routledge.

Gallese, V. (2005). 'Being like me': Self-other identity, mirror neurons and empathy. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From cognitive neuroscience to social science: Mechanisms of imitation and imitation in animals* (pp. 101–118). Cambridge: MIT Press.

Garzón, F. C. (2008). Towards a General Theory of Antirepresentationalism. *The British Journal for the Philosophy of Science, 59*, 259–292.

Godfrey-Smith, P. (2005). Folk psychology as a model. *Philosopher's Imprint, 5*, 1–16.

Goldman, A. (1989). Interpretation psychologized. *Mind & Language, 4*, 161–185.

Goldman, A. (2006). *Simulating minds: The philosophy, psychology and neuroscience of mindreading*. New York: Oxford University Press.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge: MIT Press.

Gordon, R. M. (1986). Folk psychology as simulation. *Mind and Language, 1*, 158–171.

Gordon, R. M. (2008). Beyond mindreading. *Philosophical Explorations, 11*(3), 219–222.

Hala, S., Hug, S., & Henderson, A. (2003). Executive functioning and false-belief understanding in preschool children: Two tasks are harder than one. *Journal of Cognitive Development, 4*, 275–298.

Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science, 6*, 346–359.

Heal, J. (1986). Replication and Functionalism. In J. Butterfield (Ed.), *Language, mind, and logic*. Cambridge: Cambridge University Press.

Herschbach, M. (2008). Folk psychological and phenomenological accounts of social perception. *Philosophical Explorations, 11*, 223–235.

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology, 117*, 500–544.

Hurley, S. (2008). The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *The Behavioral and Brain Sciences, 31*, 1–58.

Hutto, D. D. (2004). The limits of spectatorial folk psychology. *Mind and Language, 19*, 548–573.

Hutto, D. D. (2008). *Folk psychological narratives: The sociocultural basis of understanding reasons*. Cambridge: MIT Press.

Hutto, D. (2011). Elementary mind-minding enactivist style. In A. Seemann (Ed.)., *Developments in Philosophy, Psychology, and Neuroscience*. Cambridge: MIT Press.

Izhikevich, E. M. (2007). *Dynamical Systems in neuroscience: The geometry of excitability and bursting*. Cambridge: MIT Press.

Kovács, A., Teglas, E., & Endress, A. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330*, 1830–1834.

Lakoff, G., & Johnston, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind". *Trends in Cognitive Sciences, 8*, 528–533.

Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology, 50*, 45–85.

Lohmann, H., & Tomasello, M. (2003). The role of language in the development in false belief understanding: A training study. *Child Development, 74*, 1130–1144.

Maibom, H. (2003). The mindreader and the scientist. *Mind and Language, 18*, 296–315.

Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. Heidelberg: Springer.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*, 838–850.

Meltzoff, A. N., & Brooks, R. (2001). 'Like me' as a building block for understanding other minds: Bodily acts, attention, and intention. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 171–191). Cambridge: MIT Press.

Mitchell, P. (1996). *Acquiring a conception of mind. A review of psychological research and theory*. Hove, E. Sussex: Psychology Press.

Moore, C., Pure, K., & Furrow, D. (1990). Children's understanding of the modal expression of speakers certainty and uncertainty and its relation to the development of representational theory of mind. *Child Development, 61*, 722–730.

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

Noë, A. (2004). *Action in perception*. Cambridge: MIT Press.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *The Behavioral and Brain Sciences, 24*, 939–973.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255–258.

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences, 3*(9), 337–344.

Perner, J., Lang, B., & Kloo, D. (2002). Theory of Mind and Self Control: More than a common problem of inhibition. *Child Development, 73*, 752–767.

Perner, J., & Ruffman, T. (2005). Infant's insight into the mind: How deep? *Science, 308*, 214–216.

Port, R. F., & Van Gelder, T. (1995). *Mind as motion: Explorations in the dynamics of cognition*. Cambridge: MIT Press.

Ratcliffe, M. (2006). 'Folk psychology' is not folk psychology. *Phenomenology and the Cognitive Sciences, 5*, 31–52.

Ratcliffe, M. (2007). *Rethinking commonsense psychology: A critique of folk psychology, theory of mind and simulation*. Basingstoke: Palgrave Macmillan.

Robbins, P., & Aydede, M. (2009). *A short primer on situated cognition* (pp. 3–10). Cambridge: Cambridge University Press.

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does Eye gaze indicate knowledge of false belief: Charting transitions in knowledge. *Journal of Experimental Child Psychology, 80*, 201–224.

Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? *Trends in Cognitive Sciences, 9* (10), 462–463.

Rupert, R. (2009). *Cognitive systems and the extended mind*. New York: Oxford University Press.

Russell, J. (1996). *Agency: Its role in mental development*. Hove: Erlbaum.

Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about identity at 18 months. *Child Development, 80*, 1172–1196.

Shannon, B. (1993). *The representational and the presentational*. London: Harvester Wheatsheaf.

Sirois, S., & Jackson, I. (2007). Social cognition in infancy: A critical review of research on higher order abilities. *The European Journal of Developmental Psychology, 4*, 46–64.

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Science, 7*, 343–348.

Sodian, B. (2005). Theory of mind. The case for conceptual development. In W. Schneider, R. Schumann-Hengsteler, & B. Sodian (Eds.), *Young children's cognitive development. Interrelationships among working memory, theory of mind, and executive functions* (pp. 95–130). Hillsdale: Erlbaum.

Sodian, B. (2010). Theory of mind in infancy. *Child Development Perspectives 4*(3), 267–271.

Song, H., & Baillargeon, R. (2008). Infants' reasoning about others' false perceptions. *Developmental Psychology, 44*(6), 1789–1795.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by two-year-olds. *Psychological Science, 18*, 587–592.

Stewart, J., Gapenne, O., & Di Paolo, E. A. (2011). *Enaction: Towards a new paradigm for cognitive science*. Cambridge: MIT Press.

Spaulding, S. (2010). Embodied cognition and mindreading. *Mind & Language, 25*(1), 119–140.

Spivey, M. (2007). *The continuity of mind*. New York: Oxford University Press.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science, 18*(7), 580–586.

Thompson, E., & Stapleton, M. (2009). Making sense of sense-making: Reflections on enactive and extended mind theories. *Topoi, 28*, 23–30.

Thompson, E. (2005). Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the Cognitive Sciences, 4*, 407–427.

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Cambridge: Harvard University Press.

Torrance, S. (2006). In search of the enactive: Introduction to special issue on enactive experience. *Phenomenology and the Cognitive Sciences, 4*, 357–368.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge: Harvard University Press.

Träuble, B., Marinovic, V., & Pauen, S. (2010). Early theory of mind competencies—Do infants understand others' beliefs? *Infancy, 15*, 434–444.

Van Gelder, T. (1995). What might cognition be, if not computation? *Journal of Philosophy, 92*, 345–381.

Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *The Behavioral and Brain Sciences, 21*, 615–665.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge: MIT Press.

Walter, S. (2010). Locked-in syndrome, BCI, and a confusion about embodied, embedded, extended, and enacted cognition. *Neuroethics, 3*, 61–72.

Watson, A., Painter, J., & Bornstein, M. (2002). Longitudinal relations between 2-year-olds' language and 4-year-olds' theory of mind. *Journal of Cognition and Development, 2*, 449–457.

Weber, A., & Varela, F. J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences, 1*, 97–125.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of Theory of Mind development: The truth about false-belief. *Child Development, 72*(3), 655–684.

Wellman, H. M. (2002). Understanding the psychological world: Developing a theory of mind. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 167–187). Oxford: Blackwell.

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