Title
Embodied Representations: What are the Issues?

Permalink
https://escholarship.org/uc/item/6vm3x1jf

Journal

ISSN
1069-7977

Authors
Svensson, Henrik
Ziemke, Tom

Publication Date
2005

Peer reviewed
Abstract
A common theme of many discussions among embodied cognition researchers is whether or not cognition is making use of representations. This debate has recently been called a “debate for the sake of appearance” because there is no agreed upon way of identifying what is a representation. Therefore, instead of elaborating our own position on whether embodied cognition uses representation, we structure the debate by (1) reviewing three types of mechanisms in models of embodied cognition that might be considered candidates for representationhood and (2) outlining criteria for what is or is not a representation in embodied cognition.

Keywords: embodied cognition; interaction; representation

Introduction
With the advent of theories of embodied cognition, in the 1980s, the traditional notion of representation has been attacked by many researchers questioning the usefulness of representational explanations in cognitive science (e.g., Brooks, 1991; Beer, 1995). For example, Brooks (1991) famously argued for “intelligence without representation” and emphasized continuous real-time interaction between an embodied agent and its environment. Others have argued that representations are necessary even in theories and models of embodied cognition (e.g., Clark & Grush, 1999; Dorffner, 1997; Grush, 2004). This has led to further confusion about the concept of representation calling for “an ongoing critical assessment of representational notions” (Beer, 2003, p. 237). Haselager, de Groot, and van Rappard (2003) have suggested that the debate between representationalists and anti-representationalists is in fact futile because of the ubiquitous and often unconstrained use of representations in explanations of cognition. Haselager et al. did not commit themselves on the issue of whether cognition is representational or not. Instead, they argued that as long as there is no way of identifying actual representations in cognitive systems to constrain the notion of representation, “the debate between representationalists and anti-representationalists is bound to remain a debate for the sake of appearance” (p. 21; emphasis added; cf. also Beer, 2003; Wheeler, 2001). The purpose of this paper is to continue in the same spirit as Haselager et al., i.e., instead of elaborating our own position on whether or not cognition uses representation, we would like to contribute to structuring the debate. In order to resolve the confusion, we believe it necessary to: (1) review what types of mechanisms, processes, states, etc. can be identified as candidates for representationhood and (2) explicate the relevant theoretical frameworks and criteria for what is or is not an embodied representation.

The Representation Debate: What is the Problem?
As seen in the recent target articles by Beer (2003) and Grush (2004) as well as the respective peer commentary (e.g., Clark, 2003; Edelman, 2003; Stojanov & Bickhard 2004; Yang & Wang, 2004), the discussion is often framed (or perceived) as a debate between anti-representationalists emphasizing continuous interaction with an environment, and representationalists arguing that representations are needed for scaling up to higher-level cognition. Beer (2003) noted in his response that “despite the enormous explanatory weight that the notion of internal representation is required to bear in cognitive science, there seems to be very little agreement about what internal representations actually are” (p. 304). This was illustrated by the fact that different commentators described Beer’s example model as both representational and non-representational. Similarly, Grush’s (2004) recent emulation theory of representation
(cf. Clark & Grush, 1999) has striking similarities to Hesslow’s (2002) simulation hypothesis, but nevertheless the former is explicitly framed as a theory of representation, where the latter is decidedly anti-representational. Therefore, in this section we outline some possible reasons for this confusion and argue that these problems might be viewed differently in embodied cognition.

A possible reason for this type of confusion is according to Haselager et al. (2003) that the current definitions of representation are not capable of identifying physical states that are representational and separating them from other non-representational physical states. In contrast to many others, Haselager et al. explicitly argued for a realist explanation of representation (as opposed to an instrumental one; cf. Kim, 1998). That means, to have fruitful debates about the representational nature of cognition one needs to find “an identifiable physical state within a system that stands in for another (internal or external) state and that as such plays a causal role in (or is used by) the system generating its behavior” (Haselager et al., 2003, p. 8). The representations, according to this view, need to be not mere theoretical entities used to predict empirical data, but actual physical (neural) states (cf. Wheeler, 2001).

Haselager et al. (2003) argued that isomorphism, the standard method to detect representation, is not sufficient for separating physical representational states from non-representational ones. Isomorphism identifies representations “as those physical states that correspond one-to-one with the content carrying states outlined on the computational level and that have the causal connections to match the computational transformations” (Haselager et al., 2003, p. 8). This mapping between the two levels is a fundamental prerequisite in traditional cognitivist cognitive science. Gardner (1987), for example, argued that “it is necessary to speak about mental representations and to posit a level of analysis wholly separate from the biological or neurological” (p. 6). That means, the explanatory work is not done on the physical level but only on the higher representational level (see e.g., Fodor, 1975; Fodor & Pylyshyn, 1988).

Research in embodied cognition, on the other hand, often does not start with psychological states but with sensorimotor interaction, i.e., physical state transitions typically in more or less direct (reactive) interaction with the world making them more concerned with the physical mechanisms behind cognition. The problem with starting out at a non-cognitive level is that it makes isomorphism, as described above, insufficient for separating representational and non-representational states in physical systems (Haselager et al., 2003). However, embodied cognition could also entail new ways of conceptualizing and identifying representations. The next section provides the first steps to identifying such new constraints and conceptualizations of representations by presenting some candidates for representationhood within embodied cognition.

### Embodied Representation: Which are the Candidates?

In this section we review what we think are the main types of mechanisms, processes, or states in current models/theories of embodied cognition that can be considered as candidates for representationhood: interactive potentials, simulation or emulation, and externalization. However, the descriptions are, due to the limited space, fairly brief.

#### Interactive Potentials

Bickhard (1998; 2000; Bickhard & Terveen, 1995) has argued that representation emerges from the interactive nature of biological systems, i.e., the fact that for embodied creatures, actions influence subsequent input processing in a direct way. The interactive nature of embodied organisms leads to the idea that an agent can represent because of processes based on indications of interactive potentials and the anticipation or indication of interaction outcomes (Bickhard, 1998).

This type of mechanism is called for in situations where an organism has several interactive options available and a selection is necessary, such as the situation of the frog in the following example.

A frog seeing a fly might set up indications of the possibility of tongue-flicking-and-eating, while a frog seeing a shadow of a hawk might set up indications of the possibility of jumping in the water. A frog seeing both needs some way to decide, and internal outcome indications provide a basis for such a decision (Bickhard, 1998, p. 59).

Thus, indications of interaction outcomes have the ability to differentiate different types of environments, but it is important to notice that the indications do not necessarily correspond to any particular environment. They are only internal, functional specifications, which enable an implicit content about the agent’s presuppositions of the “appropriate” properties (Bickhard & Terveen, 1995). Interactive potentials are constituted as anticipations, which have the advantage of being true or false for the system itself. According to Bickhard (1998), it is only through interactive potentials based on anticipation or indicated outcomes that the agent itself is able to detect an error. The error is detectable because the anticipation of internal outcomes can be made false by executing the interactive potential that pointed to the outcome in an environment that does not support the type of interaction that would lead to that internal outcome.

---

3 Isomorphisms between internal and external states might be equally problematic for theories of representation within (embodied) cognition (cf. e.g. Bickhard & Terveen, 1995). However, since internal-external isomorphism is not part of any of the candidates and criteria described in later sections, we do not discuss that possibility.
Emulation & Simulation
Another recent effort to explain higher-level cognition and representation comes from so-called emulation or simulation theories (e.g., Grush, 2004; Hesslow, 2002) which are kinds of embodiment theories that have emerged in different contexts (for a review see Svensson & Ziemke, 2004; Svensson, Lindblom, & Ziemke, in press).

The basic idea is that neural structures that are responsible for perception and action, i.e. direct interaction with the world, are also reactivated and used in various cognitive tasks, such as perception, imagery, reasoning, and language. Interestingly, this basic idea is very similar in Hesslow’s (2002) (anti-representational) simulation hypothesis and Grush’s (2004) emulation theory of representation, and, to some extent, in Barsalou’s (1999) perceptual symbol systems.

Simulation or Emulation Mechanisms The main mechanisms involved in simulation is the reactivation of sensorimotor states in the absence of overt activity and the ability to anticipate consequences of that sensorimotor activity if actually performed overtly (cf. e.g., Clark & Grush, 1999; Grush, 2004; Hesslow, 2002). Hesslow (2002) summarized his simulation theory in three basic mechanisms:

1. Simulation of actions: we can activate motor structures of the brain in a way that resembles activity during a normal action but does not cause any overt movement. (2) Simulation of perception: imagining perceiving something is essentially the same as actually perceiving it, only the perceptual activity is generated by the brain itself rather than by external stimuli. (3) Anticipation: there exist associative mechanisms that enable both behavioral and perceptual activity to elicit perceptual activity in the sensory areas of the brain. (Hesslow, 2002, p. 242)

That means, internal simulations of perception and action can be coupled through an anticipatory mechanism, which enables the agent to simulate chains of behavior that range several time steps into the future (Grush, 2004; Hesslow, 2002). Essentially, simulating chains of behavior is the process by which (simulated) perceptual activity generated by a simulated action can serve as a stimulus for a new response, either in the form of an overt action, a covert simulated action or new perceptual activity.

Barsalou (1999; cf. Barsalou et al. 2003) proposed a similar kind of simulation mechanism based on perceptual symbols, which are records of the neural states that underlie perception. That is, during perception, the sensorimotor activity in the brain is recorded and thus captures information about the perceived events. The perceptual symbol is not a symbol in the sense that it by itself stands for something in the world, but can only be construed as a representation in that sense when it is used by a mechanism called a simulator. The simulator enables objects and events to be connected enabling the agent to, e.g., simulate event sequences such as what would happen if you put your hand on a hot plate. Except for the simulation and emulation accounts’ emphasis on motor processes, perceptual symbol systems can be said to be roughly based on the same general mechanism in which reactivated perceptual and/or motor states are coupled into simulated chains of behavior.

Externalization
In contrast to traditional cognitive science, some theories of embodied cognition have suggested that cognition often reaches out into the world allowing external entities to act as representations in the cognitive system consisting of both individual and environment (e.g., Clark & Chalmers, 1998; Hutchins, 1995; Zhang, 1997). The perhaps most obvious use of this kind of off-loading or externalization mechanism is external reminders. Vygotsky (1929/1977), already in the 1920s, emphasized this type of activity. For example, he used the example of tying a knot in a handkerchief as a reminder, which means that the person is “constructing the process of memorizing by externally forcing an external object to remind him of something” (Vygotsky, 1929/1977, p. 68). Since then, several other examples of ways in which humans and animals exploit external states have been reported in the literature (e.g., Larkin & Simon, 1987; Reisberg, 1987; Zhang, 1997; Susi, in press).

Studies of external representation are often not explicitly interested in theoretical discussions of what is or what is not an actual representation, but rather, how these external entities affect the individual cognizer’s ability to perform a certain task and provide a more genuine understanding of the distributed nature of cognition. Clark and Chalmers (1998), on the other hand, explicitly argued that external entities if coupled with a cognitive system can be regarded as cognitive states with the same status as other cognitive states. Even though they do not explicitly mention representations, it follows that external representations could be considered mental representations. However, this does not mean that they are arguing that something non-living can have mentality (cf. Susi, Lindblom, & Ziemke, 2003).

Embodied Representation: Which are the Criteria?
So far a number of possible representational mechanisms within models/theories of embodied cognition have been identified and briefly described. This section explicates some theoretical frameworks and criteria that might be useful in answering the question whether or not these candidate mechanisms should be considered representations. The section proceeds by outlining the criteria of representationhood that are found in five theories of representation in embodied cognition. However, the point is not to determine what will turn out to be the correct criteria to identify actual representations but merely to point out the underlying criteria for representationhood in embodied cognition.
(Radical) Constructivism & Enactivism

The radical constructivism framework (e.g. von Glaserfeld, 1995) and the enactive cognition framework (e.g. Maturana & Varela, 1987; Varela, Thompson & Rosch, 1991) emphasize cognition and knowledge as the active construction of a subject, rather than passive re-presentation of an external reality. From a radical constructivist point of view, the notion of representation as standing in is nonsensical because radical constructivism normally does not view cognition as the manipulation of representations of an independent reality, but the reality only “perturbs” (see e.g. Maturana & Varela, 1987, p. 98-99) the semantically closed and self-organizing cognitive system (Peschl & Riegler, 1999). In this sense, there are no representations but only presentations (von Glaserfeld, 1995) in that the ontological reality never enters into the representational function and gets to be re-presented in the cognitive agent (Peschl & Riegler, 1999). To consider anything a representation according to this theoretical framework the traditional notion of representation as a stand in for things in the world has to be abandoned in favor of representation as internal states or processes constructed by the individual subject, which have the power to achieve adaptive behavior and thus might not be generally thought of as re-presentations.

Interactivist Representations

Bickhard (e.g. 1998, 2000; Bickhard & Terveen, 1995) has constructed a theory of representation, interactivism, based on the pragmatist school of thought, which has also influenced various constructivist approaches including radical constructivism.

According to interactivism, any theory of representation must be capable of explaining the ability of representations to have a truth value and not only truth value per se. The agent itself must have the possibility to detect error or misrepresentation. Bickhard (2000; Bickhard & Terveen, 1995) argued that this means that representation cannot be based on any notion of correspondence or isomorphism to establish representational content, because such accounts find themselves trapped by the fact that any check for truth value must be achieved by comparing representations to the things and events in the external world. The problem with this is that the only epistemic access the agent has to the world is through those representations, thus any use of representation to check representation will be circular (e.g. Bickhard, 2000).

It is also evident that it is necessary for any representation to be embedded within an agent capable of interaction and consequently setting up expectations of future interactions (Bickhard, 1998), because representational content is only achieved by connecting earlier with further potential outcomes, where the former has been tested through interaction (Bickhard, 2000). To be more precise, according to Bickhard (2000), an indicated potential outcome has no explicit representational content. It has only implicit representational content in that it only differentiates different environments, which means that there is only an informational relationship not an explicit representation that says anything about the properties of the environment. Only when the indicated future outcomes are used to create future indications, which may e.g. hold when all “X-type environments are indicated to be also Y-type environments” (Bickhard, 2000, p. 70).

Guidance Theory

Recently, Rosenberg & Anderson (2004, forthcoming) have developed a theory of representation, guidance theory, based on the assumption that action and interaction is essential for any cognitive being. According to guidance theory, for a token to be a representation it is not important whether the representation co-varies with, is isomorphic to, or is in causal connection to the entity it tracks. In other words, what the representation is is not the right question to ask, but instead any theory of representation need to ask what the representation does for the cognitive subject. The crucial insight is that it is the cognitive subject use of the token to guide, i.e. determine, which particular motor or cognitive action it will perform that makes it a representational token (Rosenberg and Anderson, 2004, forthcoming).

Thus, guidance theory does not specify any particular form or any particular kind of correspondence with the external world for a token to be a representation. What matters is that the token are used in a particular circumstance (the subject’s internal state and immediate environment) to guide actions. As such, it defines representation as something the agent does to achieve a particular state. This means that the representations are thus closely linked to expectations of what state should be achieved (Rosenberg & Anderson, forthcoming), which is also emphasized by interactivism.

Emulation Theory & Simulation Hypothesis

Grush (2004) constructed his theory of representation on the control theoretic concept of forward models or emulators, which similar to interactivism and guidance theory emphasizes expectations or predications of future states. Representations, according to emulation theory, are found in emulator circuits in the brains of cognitive beings. Emulation circuits calculate a forward mapping from control signals, i.e., motor commands to the anticipated (sensory) consequences of executing that specific control command. However, the crucial fact of representations is that they stand in and that they can be used by the cognitive system as stand ins. That is, a representation is an internal state that can be used, in this case by emulator mechanisms, to stand in for extra-neural states of affairs when these are not available to the agent by any other means except through the stand ins, as opposed to direct causal connection with something in the agent’s near environment (cf. Clark & Grush, 1999). Representation, thus, requires a kind of decoupled internal states, which can reliably track the
normal course of affairs that the cognitive agent pursues in continuous interaction with its environment.”

Furthermore, emulation theory identifies representation as involving actual sensorimotor states and predictions of sensorimotor activations in the brain (cf. Keijzer, 2003).

As previously pointed out, while they posit similar mechanisms, the simulation hypothesis (Hesslow, 2002) is decidedly non-representational, the emulation theory of representation and perceptual symbol systems (Barsalou, 1999) keeps the representational vocabulary. Although this is not a crucial matter, Hesslow probably did not reject representations as decouplable inner states, but rather “extravagant ontological assumptions…about the existence of ‘images’, ‘representations’ or other mental entities” (p. 246); not the ability of simulation mechanisms to be used as stand ins in that sense.

External Representation
Distributed cognition and external cognition approaches extends the view of representation to include also external artifacts as representations. Representations are not only entities inside the head of a cognitive agent that are supposed to correspond to objects and events in the external world. Rather in the case of external representations, the representational states not only represent things in the sense that they correspond to something else or are about something. Rather, external representations can be identified by the different purposes they serve for cognitive agents (cf. Susi, in press). The external states or are not only inputs or representations of other entities, but are often able to affect cognition directly (e.g., Zhang, 1997). In other words, the “representation can determine what information can be perceived, what processes can be activated, and what structures can be discovered from the specific representation” (Zhang, 1997, p. 213).

External representations are external entities that directly guide actions and cognition, without being transduced into any kind of internal language of thought. What makes these structures representations is their effect on the cognitive agent, but they do not contain content independently of a cognitive observer. In this respect, external representation is not interactive representation, which identifies representationhood as anticipatory processes inside the interactive agents. However, external representations often figure as theoretical entities (cf. Grush, 2004), rather then in the strictly realist sense called for by Haselager et al. (2003).

Summary & Conclusion
This paper has outlined a number of candidates and criteria for cognition not based on cognitivist representations identified through one-to-one mappings (or correspondences) between computational structures and physical structures. Instead, the view of representationhood that emerges from the field of embodied cognition is rather different as it does not focus that much on the particular structure or form of the representations but rather their use and origin in interaction able cognitive agents. In general, the embodied cognition candidates and criteria for representation differ from the isomorphism criterion because the representations are not identified as psychological or computational states. The criteria for being a representation stems instead from what the representation does for the agent or more specific the agents interactions with the world, it is no longer representation as an internal mirror of an observer-independent external world that is motivating the use of representations in cognition.

Although the analysis is admittedly somewhat preliminary, we will briefly explicate some of the relations between the candidates and criteria for representationhood. According to radical constructivism, all of the posited mechanisms can be considered to create adaptive behavior for an agent, but as pointed out above this might not be a criterion for re-presentation per se. Interactivist representations must be able to be true or false for the agent. Consequently, representations are internal to an agent, making external representations less likely candidates for representationhood. Whether or not emulation also accounts for the emergence of interactive representation is an open issue (cf. Grush, 2004; Stojanov & Bickhard, 2004). Guidance theory shares a number of assumptions about representation with interactivist representation, such as truth value for the agent and emphasis on action and interaction. However, the theory could perhaps accept external representations given that they can provide guidance to a cognitive agent with respect to what it represents (cf. Rosenberg & Anderson, forthcoming). Emulation and simulation theory emphasize the presence of decouplable states, which could include both interactive potentials and external representations (cf. Grush, 2004; Dartnall, 2004). External representation is not mainly about meaning, but could perhaps be made part of such an enterprise if considering the use of external entities in agent-environment interactions (cf. Dartnall, 2004).

To conclude, the purpose of this paper was not to argue for or against any notion of representation nor determine what role any of these representations have in cognition. Instead, we wanted to help structure the debate by contrasting some of the possible alternatives and (sometimes hidden) assumptions behind discussions of representations in embodied cognition. In this respect, the paper has identified three candidates for representationhood, and criteria in models/theories of embodied cognition that can be used to constrain notions of representations, and also help others see the assumptions behind some of the representational mechanisms proposed in embodied cognition. This, we hope, can help to clarify discussions about representations in cognitive science.

Acknowledgements
We would like to thank Jessica Lindblom and Tarja Susi for comments and discussions contributing to this paper.

---

4 See also discussions in Wheeler (2001) and Haselager et al. (2003).
References


