
VALENTINA CUCCIO

Università degli Studi di Palermo

valentina.cuccio@unipa.it

THE NOTION OF REPRESENTATION AND THE BRAIN

abstract

The definition of the mechanism of Embodied Simulation is controversial. To account for this mechanism, Goldman and de Vignemont 2009 proposed the notion of mental representation in bodily format. In this paper I will offer arguments against the definition of mental representation in bodily format. To this purpose, I will specifically focus on the distinction between personal and subpersonal levels of explanation and on a first-person approach to the study of mental phenomena.

keywords

Embodied simulation, mental representation, first-person approach

1. Introduction

The term *representation* is, perhaps, one of the most contested expressions in the history of philosophy. Hundreds of pages would not suffice to summarize all of the different definitions and usages it has undergone in centuries of philosophical inquiry. Thus, my aim here must be far more limited. I will focus on some usages of this notion in the field of embodied cognition. Even in embodied cognition studies, the notion of representation has been seen in very different lights. Radical enactivists claim that we should get rid of this notion altogether, while many other supporters of embodiment believe that how the mind works cannot be explained without it. In fact, the use of the notion of representation in theories about human cognition has been considered as a demarcation line between radical and less radical embodied theorists (this distinction was introduced by Clark 1997; see also Chemero 2009; Alsmith & de Vignemont 2012). On one hand, radical embodied theorists claim that we can explain how the human mind works without resorting to mental representations (Kelso 1995; Port & van Gelder 1995; Thelen & Smith 1994; van Gelder 1995; Chemero 2009); on the other hand, exponents of moderate embodiment propose theories that include both representational and not representational explanations of human cognition (Barsalou 1999, 2008; Clark, 1997; Gallese & Sinigaglia 2011; Goldman & de Vignemont 2009). I will not analyze all of the usages representation has undergone in embodiment literature here. Instead, I will focus on a more specific problem: to what extent can we define the mechanism of Embodied Simulation in terms of mental representations? Embodied Simulation is the activation of specific neural circuits of the brain that control actions, perceptions, the experiencing of bodily states or emotions when a person is not actively engaged in those actions, perceptions, bodily states or emotions. To give an example, the motor areas in my brain that control the action of grasping a cup will be activated not only when I effectively grasp a cup but also when I see a cup, when I observe someone else grasping a cup, when I read or listen to a proposition about someone grasping a cup or when I just imagine someone grasping a cup. Today, the question about the legitimacy of the definition of the mechanism of Embodied Simulation in terms of mental representation is very urgent. The characterization of this mechanism in the current debate is quite controversial and, as a consequence, the definition of its role in human cognition is also quite controversial (see Mahon & Caramazza 2008 for a deeper discussion of this point).

In a recent paper, Goldman and de Vignemont (2009) proposed the notion of mental representations in bodily format. In the authors' words, these mental representations are identified with the activation of the mirror mechanism that gives rise to Embodied Simulation. This is clearly stated at the very beginning of their paper.

We offer several interpretations of embodiment, the most interesting being the thesis that mental representations in bodily formats (B-formats) have an important role in cognition. Potential B-formats include motoric, somatosensory, affective and interoceptive formats. The literature on mirroring and related phenomena provides support for a limited-scope version of embodied social cognition under the B-format interpretation. (Goldman and de Vignemont 2009, p. 154).

The notion of mental representation in bodily format was later explicitly adopted by Gallese and Sinigaglia (2011) and a similar concept can also be found in Barsalou's idea of grounded symbols (Barsalou 2008). To what extent can we legitimately define Embodied Simulation in terms of mental representations? For the mechanism of simulation being considered as a mental representation, it would be necessary that we could clearly distinguish between the content and the format of the representation. Furthermore, it would also be necessary to identify the subject of the mental representation. I will suggest that neither of these criteria is matched by the notion of mental representation in bodily format. To this purpose, I will discuss some problematic issues related to the definition of this notion and I will specifically focus on the distinction between personal and subpersonal levels of explanation and on a first-person approach (Baker 2013) to the study of mental phenomena.

2. The Notion of Representation in Brain Sciences

The term *representation* has been widely used in brain sciences. In many cases, the use of this term does not imply any content-bearing state. To give an example, the use of this term has been fairly common in the somatotopic description of the brain. Somatotopy is the identification of the correspondences between areas of the brain and parts of the body. Somatotopy tells us which part of the brain *represents*, namely *controls*, movements of the face, hands, legs and so on. These causal mappings between parts of the brain and certain areas of our body have been described and explained by means of an image, the Penfield *homunculus*, drawn in the human cortex. Although characteristics of the Penfield *homunculus* have been widely questioned, this terminology is still present in

the debate. The *homunculus* drawn in the human cortex and its description in representational terms are just a visual metaphor of the correspondences between areas of the brain and parts of the body.

Cognitive neuroscientists also use the term representation when they refer to our motor repertoire. It is common, in this case, to say that neurons represent the goals of actions such as grasping or kicking. It is also common to talk about neurons that “represent” human faces or objects and so on. How can we interpret the use of the term representation in those cases? Are neuroscientists endorsing an intentionalist and representationalist account of what neurons do? Let us look closely at the case of motor neurons representing goals. A deeper analysis of this case will help us to show that the use of the term representation or of other expressions such as goals or intentions, that seem to ascribe mentalist and representationalist power to neural circuits, does not really imply a mentalist or representationalist explanation.

Neuroscientific evidence tells us that the motor cortex has a goal-centered organization (Umiltà et al. 2008). That is, there are neurons in the motor cortex that represent the goals of actions, independently of the specific movements we accomplish to carry out those actions. For example, Umiltà and colleagues (2008) showed that the “grasping” neurons in the F5 area of the motor cortex fire both when a monkey grasps an object by using normal pliers as well as when it uses inverse pliers. Normal and inverse pliers require hand-movement-patterns that are opposite from one another. Physiology principles of correlational learning can explain how we build our motor repertoire and the goal-centered organization of the motor cortex. We know that neurons that fire together for a sufficient amount of time start to strengthen their mutual connection, thus creating a neural circuit. This principle is known as long-term potentiation (see Hebb 1949). By means of long-term potentiation and other phenomena of correlational learning, different areas of the motor cortex become wired together. That is, they are physically wired in a chain. Thus, neurons that usually fire in the F5 area of the monkey’s brain when the monkey grasps an object with her fingers can also fire when the monkey grasps the object with normal or even inverse pliers. But to have this result, it is necessary to train the monkey. That is, it is necessary to create that physical chain of neurons. Umiltà and colleagues (2008, 2011) explain how this mechanism works:

What could be the mechanism that allows a transformation of a goal into appropriate movements even when an opposite sequence of movements is necessary to achieve the goal? Our findings show that, after learning, the

correct movement selection occurred immediately as soon as the monkey grasped one or the other type of pliers. This correct movement selection may be accounted for if one admits that goal-related F5 and F1g neurons are synaptically connected with two different sets of motor cortex neurons controlling the opening and the closing of the hand, respectively. These movement-related neurons, besides sending their output to the spinal cord, would also send a corollary discharge to the goal-related F5 and F1g neurons. In a natural setting, daily interactions with objects reinforce the connections that lead to the desired goal, thus selecting first those neurons that control hand opening and then those that control hand closure. After learning to use the reverse pliers, the opposite connections, reinforced by the success of the tool-mediated motor acts, prevail. As a consequence, the neurons that control hand closure are selected first, and those that control hand opening are selected subsequently (Umiltà et al. 2008, p. 2211).

It is clear that, in the case of neurons representing goals, we are describing physical chains of neurons that are synaptically connected and that are the result of correlational learning. In other words, this is an entirely mechanical and physical process. No *goal* is *represented* by those neurons that is in any sense different from a mechanical description. Similar results have also been observed in a study that recorded Motor Evoked Potentials to TMS from the right *opponens pollicis* of humans when using normal or reverse pliers or when observing others using the same tools, both with a specific goal (to grasp something) or without any specific goal (Cattaneo et al. 2009). According to the authors' interpretation of the data, the goal-centered organization of the motor cortex allows us to understand other people's goal-oriented actions by means of a kind of generalization. The observation of any grasping action, independently of the specific movement involved in the action, for example even when we are observing the use of tools such as inverse pliers, makes our implicit knowledge of real grasping available to us. Furthermore, according to the authors' interpretation, we can say that when we observe goal-oriented actions carried out by means of tools, the tool is incorporated in the observer's body-schema. This clearly explains why, when we observe people using reverse pliers with a goal, "grasping" neurons in the F5 motor cortex are activated by a pattern of movements that are the opposite of those of the real grasp or of a standard pliers grasp. In this case, reverse pliers become the distal effector that determines the activation of grasping neurons. Should we endorse an intentionalist and representationalist interpretation

of what neurons do? The fact that the same “grasping” neurons in the F5 motor cortex control flexor and extensor muscles being synaptically connected to both of them, as was suggested by Umiltà et al. (2008), and the hypothesis advanced by Cattaneo et al. (2009) about the incorporation of tools in the observer’s body schema, which explained why “grasping” muscles are activated by the observation of grasping with reverse pliers that involve an opposite pattern of movement, seem to suggest that the use of the term representation in this case does not allow for a representationalist interpretation. In the first case, we have a physical chain of neurons that does not necessarily imply any representational relationship. In the second case, reverse pliers become a distal effector. The reverse pliers movement resembles the movement of a real grasping action and, when we observe or execute goal-oriented actions, pliers become our own fingers. Thus, when we observe someone else using reverse pliers to grasp something, the grasping neurons in the F5 motor cortex are activated because of the mechanism of simulation. As an alternative, we could describe these cases in terms of causal relations between physical events which, in turn, function as content vehicles. Thus, the goal-related F5 neurons would be bearers of a representational content. The point is, what would be the explanatory or predictive virtues of this representationalist explanation? What could this representationalist explanation add to our account of the motor system that a physical explanation cannot provide? So far, it seems that a representationalist explanation would not add anything to the picture we can sketch in physical terms.

Thus, the specific usages of the term representation discussed so far seem to be technical acceptations, internal to the neurophysiological and neuroscientific jargon. These usages do not necessarily commit neuroscientists to a representational theory of mind. However, recently and very often this technical acceptations of the term representation, in which the subject of the representation is a brain area or a particular neural circuit, is qualified as mental. Is this a correct move? Are those putatively *mental representations* the milestones on which we can construe our theory of language, social cognition, and so on? In the following section I will analyze one particular usage of the notion of mental representation in relation to neural facts, the notion of mental representation in bodily format proposed by Goldman and de Vignemont (2009).

3. Mental Representations in Bodily Format

Goldman and de Vignemont defined mental representations in bodily formats as those realized by means of the mechanism of Embodied Simulation (see the introduction to this paper). According to their definition, the activation of the hand-related areas of my motor cortex when I am looking at someone else grasping a cup is a mental representation, encoded in a motoric format, of

the action of grasping a cup. Considering that the mechanism of simulation is not limited to motor areas of the brain but it is a widespread mechanism in our brains, in the same vein, we can have representations in somatosensory, affective or visual formats, and so on.

In a previous paper (Cuccio, submitted) I have already proposed two arguments against the definition of mental representation in bodily format. These arguments can be summarized as follows. First, we cannot define the mechanism of simulation as a representation because, in this case, it is not possible to distinguish between the content and the format of the representation. This distinction is implicitly present in our usages of the notion of representation and holds true even in very different philosophical traditions. The distinction between content and format is a necessary condition to define something as a representation and this distinction cannot be applied to the case of Embodied Simulation, where neurons firing are at the same time the format of the representation and its informational content. And, indeed, while Goldman and de Vignemont (2009, p. 155) make a clear distinction between content and format when they talk about representations with bodily content, when they define the notion of representation with bodily format, which is identified with the mechanism of simulation, they are no longer able to make such a distinction. The format of the representation is entirely identified by means of its content and the informational content cannot be truly distinguished from the format. This is evident from the examples the authors provide in their paper.

Second, it has been observed that in the case of neurons firing during the process of Embodied Simulation, we cannot clearly make a distinction between the role these neurons carry out in the circuit in which they are embedded and the putative information they should convey, since they are considered as a representation. In other words, a real occurrence of a phenomenon and the occurrence of its representation should differ, while in the case of the process of Embodied Simulation, they completely overlap (see Cuccio, submitted, for a deeper discussion of these arguments). In the next section, I will discuss another argument against the notion of mental representation in bodily format. Such argument is based on the distinction between the personal and subpersonal level of explanation (Dennett 1969) and on a first-person approach (Baker 2013) to the study of mental phenomena.

4. **Personal and Subpersonal Levels of Explanation** The use of the term *mental* in the notion of mental representation in bodily format seems to be highly ambiguous. On the one hand, unless we are committed to a strong reductionist hypothesis, to qualify a process as a *mental* process seems to suggest that we are dealing with something that happens

at the personal level, the level of the people experiencing and acting. The personal level is the level that we experience from a first person perspective. On the other hand, what Goldman and de Vignemont are referring to when they define mental representations in bodily format are the patterns of neural activation. The activation of neurons is a subpersonal physical process. We can only gain epistemological access to this kind of processes from a third-person perspective. Although subpersonal physical processes are constitutive of our experiences at the personal level (see Colombo 2013), if we describe our mental processes in the third-person perspective we will eliminate from our explanations our knowledge of these experiences in the first person perspective. Yet Goldman and de Vignemont (2009) propose their definition of mental representation in bodily format without subscribing to any form of reductionism. There seems to be a contradiction here. In fact, the authors seem to be far from proposing a redefinition of the human mind in biological terms. However, can patterns of neural activation be defined as mental processes without subscribing to reductionism? It is worth noting here that, starting from the beginning of the 80s and during the 90s, researchers working in the Computational and Representational Theory of Mind proposed the idea of subpersonal mental representations as the building blocks of our cognition (Fodor 1998). These subpersonal mental representations were symbolic units of the Language of Thought that could be creatively combined according to syntactic rules and the principle of compositionality. Human cognition, it was suggested, can be entirely expressed in a propositional format.

Although this research paradigm had a great and long-standing influence on philosophical debate, it also faced some problematic aspects, particularly concerning the very same existence of those subpersonal mental representations. The question behind this problem was how a physical and mechanical system such as the brain can acquire representational content. Different solutions have been proposed to answer this question (e.g. Dretske 1981; Millikan 1984). However, all the programmes of naturalization of subpersonal mental representations had deep problems to solve such as: the problem of the indeterminacy of the content (how can a brain-state acquire a content and become a mental representation?); the problem of the decouplability of representations, which leaves very little explicative power to some of the solutions proposed; the problem of the gap between the subpersonal and the personal level of experiences; and the problem, originally proposed by Ryle, of the existence of a kind of practical, non-propositional, knowledge to ground propositionally structured knowledge. This kind of knowledge is not considered in the Representational and

Computational Theory of Mind. Hence, this issue of the very same existence of subpersonal mental representations is highly problematic and still unsolved, even in the research paradigm that originally proposed this notion. This research paradigm has currently lost momentum, also in light of the embodied theories of cognition that are currently being proposed. Many empirical findings have widely shown that the existence of amodal symbols in the mind/brain is largely unfounded (see Barsalou 2008 for a discussion).

Let us now look closer at Dennett's distinction between the personal and subpersonal level of description. Daniel Dennett introduced the distinction between the personal and subpersonal level of explanation in 1969. To define it, he referred to the case of pain.

When we ask a person why he pulled his hand away from the stove, and he replies that he did so because it hurt, or he felt pain in his hand, this looks like the beginning of an answer to a question of behavioural control, the question being how people know enough to remove their hands from things that can burn them. The natural "mental process" answer is that the person has a "sensation" which he identifies as pain, and which he is somehow able to "locate" in his fingertips, and this "prompts" him to remove his hand. (Dennett 1969, p. 91).

The personal level of explanation pertains to "the explanatory level of people and their sensations and activities". The subpersonal level, on the other hand, concerns "the level of brains and events in the nervous system" (Dennett 1969, p. 93). In other words, personal level phenomena are those *mental* processes that characterise our life as subjects, as persons, while subpersonal phenomena are physical processes that we can describe in mechanical terms. Interestingly, in relation to the pain example, Dennett says:

When we abandon mental process talk for physical processes we cannot say that the mental process analysis of *pain* is wrong, for our alternative analysis cannot be an analysis of pain at all, but rather of something else – the motion of human bodies or the organization of the nervous system (Dennett 1969, p. 93).

Thus, the personal level of explanation pertains to people. This is the level of beliefs and desires; this is the level of normative agents and moral responsibility.

Intuitively, this distinction is hard to eliminate. Lynne Baker (2013) proposes an argument against reductionism that is based on biological

considerations. The argument runs as follows. Biologists say that the differences between human and other non-human primates are biologically insignificant. On the other hand, there is a tremendous difference between human and non-human primates in cognitive, socio-cognitive and communicative terms. From both these premises, it follows that we need to go beyond biology to make sense of this difference. A similar argument is presented by Michael Tomasello (1999) to explain the evolution of mankind. In his account, a solely biological explanation would not be enough to make sense of the extraordinary fast evolutionary path that led us to be human. We need to go beyond biology and also take into consideration the cultural dimension of human evolution (on the interaction between culture and biological evolution see also Deacon 1997).

The personal level is the level of our experiences. Our entire mental life, as well as all of our various forms of mental content, would be lost if we reduced mental processes to their subphysical mechanisms (Baker 2013) because we would abandon our first-person approach to these experiences. As Dennett says, if we talk about the physiology of pain we are not talking about pain anymore, about how it feels and how we react to it (similar arguments, Dennett acknowledges, can be already found in Ryle and Wittgenstein). If we abandon this distinction, then, we will inevitably and immediately lose the legitimacy of our first-person epistemological access to the world. This is the price to pay if we do not accept Dennett's distinction. On the other hand, to accept this distinction does not necessarily lead us to endorse a dualistic approach. And, in fact, Dennett, at least in his 1969 book *Content and Consciousness*, clearly claims an anti-dualist and anti-physicalist approach to the mind-body problem. We are what we are because we have the bodies that we have. That is, even if we cannot explain our personal level experiences in subpersonal terms and even if we cannot describe the activity of a physical subsystem in mental terms that is more isolated from the rest of the system, the qualifications that we ascribe to the whole system, the person, are necessarily dependent on the body that we have. Dennett's distinction is a distinction between levels of explanation; he is not claiming the existence of two different substances.

Things are different in many respects in Dennett's later works but I am not going to address this problem here. I will just keep that distinction between the personal and subpersonal levels of description on the table, as it was formulated in his 1969 book.

If we buy Dennett's definition of these two levels of explanation, then we cannot qualify the activation of the mechanism of simulation *per se* as a mental phenomenon. We cannot do that simply because we have already

abandoned the level of people and their *mental* processes. We are talking about physical processes in the brain. Hence, as we have already seen, the only way to accept the definition of *mental* representation in bodily format and to be coherent would be to embrace a strong reductionist theory. That is, mental phenomena are physical phenomena that we can explain in terms of neural activity in the brain. Though, Goldman and de Vignemont do not seem to be committed to any reductionist hypothesis. Then, on the basis of these premises, it follows that their definition cannot be coherently accepted.

- 5. Conclusions** Having said that, what then about Embodied Simulation? Is the claim that Embodied Simulation cannot be considered as a mental representation equivalent to say that it is not necessary or not relevant in human cognition? My aim here is not to make such a claim. Embodied Simulation is a central and important mechanism in human cognition and a lot of empirical evidence supports this hypothesis. Empirical evidence suggests that Embodied Simulation is a part, at the subpersonal level, of the processes that allow us to comprehend other people actions or to understand language. Its role seems to be constitutive of the process of understanding and not merely causally correlated to it or just a side effect of the process of comprehension. In fact, when the mechanism of simulation is disrupted, for example artificially by means of TMS, the process of understanding is somehow impaired (see Pulvermüller 2013 for a discussion of the constitutive role of the mechanism of simulation during the comprehension of language).
- It is then of paramount importance to rethink the mechanism of simulation and to describe its role without appealing to the notion of mental representation.

REFERENCES

- Alsmith, A. & de Vignemont, F. (2012), "Embodying the Mind and Representing the Body", *Review of Philosophy and Psychology*, Special Issue 3(1), pp. 1-13;
- Baker, L.R. (2013), *Naturalism and the First-Person Perspective*, Oxford University Press, New York;
- Barsalou, L.W. (1999), "Perceptual Symbol Systems", *Behavioral and Brain Sciences*, 22, pp. 577-609; discussion pp. 10-60;
- Barsalou, L.W. (2008), "Grounded Cognition", *Annual Review of Psychology*, 59, pp. 617-645;
- Cattaneo, L., Caruana, F., Jezzini, A. & Rizzolatti, G. (2009), "Representation of Goal and Movements Without Overt Motor Behavior in the Human Motor Cortex: A Transcranial Magnetic Stimulation Study", *Journal of Neuroscience*, 29, pp. 11134-11138;
- Chemero, A. (2009), *Radical Embodied Cognitive Science*, MIT Press, Cambridge (MA);
- Clark, A. (1997), *Being There: Putting Brain, Body, and World Together Again*, MIT Press, Cambridge (MA), London;
- Colombo, M. (2013), "Constitutive Relevance and the Personal/Subpersonal Distinction", *Philosophical Psychology*, 26(4), pp. 547-570;
- Cuccio, V. (submitted), "Embodied simulation as bodily attitude. For a direct role of the body in language and cognition", *Philosophical Psychology*.
- Deacon, T.W. (1997), *The Symbolic Species: The Co-Evolution of Language and the Brain*, Norton, New York;
- Dennett, D.C. (1969), *Content and Consciousness*, Routledge & Kegan Paul Humanities Press, New York;
- Dretske, F. (1981), *Knowledge and the Flow of Information*, MIT Press, Cambridge (MA);
- Fodor, J. (1998), *Concepts: Where Cognitive Science Went Wrong*, Oxford University Press, New York;
- Gallese, V. & Sinigaglia, C. (2011), "What Is So Special About Embodied Simulation?", *Trends in Cognitive Sciences*, 15, pp. 512-519;
- Goldman, A., & de Vignemont, F. (2009), "Is Social Cognition Embodied?", *Trends in Cognitive Sciences*, 13 (4), pp. 154-159;
- Hebb, D.O. (1949), *The Organization of Behavior: A Neuropsychological Theory*, John Wiley, New York;
- Kelso, J.A.S. (1995), *Dynamic Patterns: The Self-Organization of Brain and Behavior*, MIT Press, Cambridge (MA), London;
- Mahon, B.Z. & Caramazza, A. (2008), "A Critical Look at the Embodied Cognition Hypothesis and a New Proposal For Grounding Conceptual

- Content”, *Journal of Physiology of Paris*, 102, pp. 59-70;
- Millikan, R. (1984), *Language, Thought and Other Biological Categories*, MIT Press, Cambridge (MA);
- Port, R.F. & van Gelder, T. (1995), *Mind as Motion: Explorations in the Dynamics of Cognition*, MIT Press, Cambridge (MA), London;
- Pulvermüller, F. (2013), “How Neurons Make Meaning: Brain Mechanisms for Embodied and Abstract-Symbolic Semantics”, *Trends in Cognitive Science*, 17, pp. 458-470;
- Thelen, E. & Smith, L.B. (1994), *A Dynamic Systems Approach to the Development of Cognition and Action*, MIT Press, Cambridge (MA), London;
- Tomasello, M. (1999), *The Cultural Origins of Human Cognition*, Harvard University Press, Cambridge (MA), London;
- Umiltà, M.A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., Jezzini, A., Gallese, V. & Rizzolatti, G. (2008), “When Pliers Become Fingers in the Monkey Motor System”, *Proceedings of the National Academy of Science USA*, 105, pp. 2209-2213;
- Van Gelder, T. (1995), “What Might Cognition Be, if Not Computation?”, *The Journal of Philosophy*, 92(7), pp. 345-381.