A New Imagery Debate: Enactive and Sensorimotor Accounts

Lucia Foglia[1] · John Kevin O'Regan[1]

Abstract Traditionally, the "Imagery Debate" has opposed two main camps: depictivism and descriptivism. This debate has essentially focused on the nature of the internal representations thought to be involved in imagery, without addressing at all the question of action. More recently, a third, "embodied" view is moving the debate into a new phase. The embodied approach focuses on the interdependence of perception, cognition and action, and in its more radical line this approach promotes the idea that perception is not a process involving internal world-models (representations). The anti-representationalist version of the embodied paradigm covers, among others that we shall not discuss here, two quite different positions, namely the enactive approach and sensorimotor theory. Up to now these two anti-representationalist accounts have generally been confounded. In this paper we will argue that despite some important commonalities, enactive and sensorimotor accounts come with distinctive theoretical traits with regard to their approach to imagery. These become manifest when critically examining the role they assign to sensorimotor engagements with the world. We shall argue that enactive and sensorimotor approaches are different in their understanding of the role of embodied action, and these different notions of embodiment lead to different explanatory accounts of perception and imagery. We propose that, due to existing ambiguities in enactivism, the sensorimotor theory is a better framework for a skill-based approach to imagery.

1 Introduction

Try to imagine an apple. You can imagine it the way you want. Maybe you imagine it to be red, maybe not. Maybe you visualize it in a fruit basket, lunch bag, or near a jar of peanut butter. You can tell us the size or brand, should we ask. You would probably say that having a visual image of an apple is similar to looking at an actual apple. Though in

✉ Lucia Foglia
foglialucia977@gmail.com

[1] Laboratoire Psychologie de la Perception, CNRS - Université Paris Descartes, Centre Biomédical des Saints Pères, 45 rue des Sts Pères, Paris 75270 Paris cedex 06, France

visual imagery there is no perceptual feedback from the world itself, hence no sensory presence, the two experiences are phenomenally and functionally alike in many respects.

Cognitive neuroscience accounts for such similarity by showing evidence that perception and imagery share neural mechanisms in the brain. Although there are not yet conclusive arguments for the common substrate, claims about some degrees of anatomical overlap are not contested. It is how to interpret this structural, neural sharing which has constituted the kernel of the imagery debate.

The "imagery debate" has traditionally opposed two views: depictivism and descriptivism. The two views dispute the nature of the representations that allegedly underlie imagery and come to different interpretations concerning evidence that vision-related processes are activated in mental visualization tasks (Pylyshyn 1981, 2002, 2003; Kosslyn 1980, Kosslyn 1994, Kosslyn 2005). Behind this debate lies the controversy on modularism, and whether clear-cut boundaries between visual processes and cognitive activity can be maintained. With the rise of the embodied paradigm (Chemero 2009; Barsalou 2008; Clark 1997, 2008; Thompson 2007; Gallagher 2005; Noë 2004; Varela et al. 1991) the debate has moved into a new phase. Within the embodied framework, which is characterized by a spectrum of different formulations, two distinct approaches have been particularly distinguished for their non-representationalist leanings [1]: enactivism (Thomas 1999, 2009, 2014; Thompson 2007) and the sensorimotor theory (O'Regan and Block 2012, O'Regan 2011; Noë 2004, 2010; O'Regan and Noë 2001). Although one might be under the impression that the position held by some advocates of enactivism is fully compatible with the view defended by sensorimotor theorists, we will argue that there are valid reasons to think that the two are not to be lumped together or used interchangeably. The aim of this paper is to show that their distinctive theoretical traits lead to different accounts of mental imagery. We will motivate our claim by drawing on their divergent views on what perceptual experience requires. Due to existing ambiguities in enactivism, we will argue in favor of the sensorimotor approach to mental imagery.

The layout of the paper is as follows. In Section 1 we aim to untangle the web of conceptual issues making up the "imagery debate" by classifying the different views with respect to their appeal to the notions of representation and action. The goal is not to put forward the explanatory advantage of theories that sidestep the issue of representation altogether, but rather to show how the debate has been affected by the turn from classic to embodied cognitive science. As we shall see, this paradigm shift has led to the formulation of a wide range of embodied views on imagery, which can be differentiated on the grounds of their being more or less conservative. Within this range, enactivism and sensorimotor theory have been usually equated for their anti-representationalist leaning. Upon closer examination, however, their association turns out to be problematic.

Section 2 lays some groundwork for Section 3. To motivate our claim that enactivism and sensorimotor theory hold two distinct accounts of imagery, we have to show first where their divergence originates. We take this divergence to stem from the fact that different interpretations of the role of action in perception are transposed into distinct views of imagery. In Section 3 we critically examine the enactivist claim

---

[1] We skip a discussion of other anti-representationalist approaches, such as Gibsonian ecological psychology, dynamical systems theory and bahaviour-based artificial intelligence, because they go beyond the scope of the present paper.

that imagery consists in mentally reenacting the perceptual exploratory behavior that would be carried out if one were actually seeing the imaged object (in short the "reenactment thesis"). We will argue that the reenactment thesis is unworkable unless it makes an appeal to representations. The sensorimotor theory is not committed to the reenactment thesis. On this view, no reconstruction, or simulation or partial recreation of perceptual acts is needed, since for mental imagery it is sufficient to be attuned to a pattern of sensorimotor regularities. Such "attunement" involves having implicit expectancies without (mentally) going through the motions. The last section contains final remarks and stresses that an appeal to sensorimotor laws explains the phenomenal similarity between perception and imagery in a plausible way.

As a terminological remark, the term imagery used throughout the text, unless otherwise specified, refers to visual imagery. Although we will not venture into discussing other types of imagery, what we shall claim about visual imagery can in principle apply to other types of imagery.[2]

2 The Imagery Debate: From Classic to Embodied Approaches

In this section we present a classification of different views on imagery. The principle of classification will concern the extent to which different theories employ the notions of action and representation. The first two theories of imagery constitute the core of what has up to now classically been called the "imagery debate". But this debate has essentially involved the question of the *nature* of the internal representations presumed to be used in imagery and the role that action plays has been either dismissed or not fully explored. We then have three intermediate or hybrid theories, which we consider rather conservative and moderately embodied because, though adding the element of action, they retain the notion of representation. Finally, we have the more radical wing of the embodied cognition spectrum constituted by situated theories such as the sensorimotor and enactive approaches. In contrast to intermediate theories, these positions drop the notion of representation altogether, and precisely for this reason they may count as full-blooded formulations of the embodied framework.

2.1 Representationalism

2.1.1 Depictivism

An influential position in the "Imagery Debate", championed primarily by Kosslyn (1980, 1994, Kosslyn et al. 2006) and his collaborators (Kosslyn et al. 1995, 1997, 1999), holds that the internal representations underlying the experience of a mental

---

[2] One more caveat: although one may be under the impression that imagery and imagination refer to the same mental phenomenon, or that imagery is a subtype of imagination, the link between the two is an open issue. Some (for example White 1990) have stressed their differences, basing their argument on the fact that various mental processes, such as memories, hallucinations, dreams, imagistic beliefs, involve imagery without involving imagination, and that whereas in imagery a subject has minimal control, imagination is under voluntary control. Though we acknowledge the relevance of the topic, it lies beyond the scope of the present paper to discuss the link and differences between the notions of mental imagery and imagination (we refer to Kind 2001 for a thoughtful investigation).

image are of the same type as those used in perception. Crucial to this position is the fact that patterns of activation in lower-level, topographically organized modality-specific brain areas (in particular, early visual areas) are used to account for the depictive nature of mental imagery. As Kosslyn (2005) acknowledges, the approach offered is reductionist: imagery essentially involves topographically organized representations, and attention shifts across such internally generated representations result in the experience of seeing with the mind's eye.

Proponents of depictivism have argued that neuroscientific evidence concerning the activation of common brain regions in perception and imagery has provided reasonable backing for this theory (Ganis et al. 2004; Kosslyn et al. 1993, 1997).[3]

2.1.2 Descriptivism

Pylyshyn (1981, 2002, 2003) is the other prominent figure of the imagery debate. His position, known as descriptivism, differs from Kosslyn's in that it offers an alternative interpretation of evidence concerning eye movements and the activation of brain systems related to vision. As for the involvement of visual cortices and what this implies for the nature of the representations used in imagery, he argues that the available collected evidence does not warrant the conclusion that mental images have a depictive format that reflects the properties of their medium of representation. Rather, he says, mental images should be taken to consist of symbolic descriptions. As for evidence that the eyes move during imagery, Pylyshyn's proposes that any behavioral output is merely an epiphenomenal byproduct, which is to be explained with reference to "tacit" knowledge consisting of implicit, not easy to verbally articulate, knowledge of the world. Consider what would happen if we were asked to imagine a large object placed next to a small one. To Pylyshyn, we would be able to answer questions about sizes without the need to assume that we are actually mentally visualizing something looming large in front of us as opposed to something being small. For a descriptivist, to imagine something large is *not* to internally visualize something with inherent spatial properties. Rather, it is to *know* that large things come with more visible details and that it usually takes longer to scan or report on something with more visible details than scanning or reporting on something with fewer ones. Put bluntly, when subjects are asked to imagine x, all they do is to draw on the knowledge that seeing x brings forth (Pylyshyn 2002). This very naturally explains why people who do not have knowledge about certain visual properties fail to report them in their imagery, and also why, for example, it takes longer to scan imaged points in conditions when a subject is told that they are further apart (Pylyshyn 1981, 2002).

2.2 Intermediate Approaches Involving Action and Representation

We now present three intermediate positions which retain the notion of representation but add the element action. What notion of representation is supposed to be at work,

---

[3] It should be noted that the present-day version of the theory includes reference to oculomotor activity (saccades, fixations and eye scanpaths) as being necessary and functional to the build-up of a depictive image (Mast and Kosslyn 2002), and accommodates a conceptual framework where mental imagery is regarded as a simulation of perceptual behavior (Moulton and Kosslyn 2009).

however, is not always entirely clear, since these positions tend to remain neutral with respect to the representationalist dispute.

Though these positions grant a non-trivial role to motor behavior, their appeal to the notion of representation makes them weaker formulations of the embodied framework when compared to more radical lines.

2.2.1 Scanpath Theory

The scanpath theory is an example of a representation-based theory which nevertheless makes the link with action by specifying how a sequence of eye fixations yields mental visualizing. According to this theory oculomotor information is stored along with information encoded at each fixation, and during imagery it is used as spatial index to correctly assemble and arrange part-images into a coherent representation (Brandt and Stark 1997; Noton and Stark 1971a, b). Evidence supporting the existence of scanpaths in mental imagery comes from experiments showing the correspondence between eye movement patterns during perceptual encoding and those observed during a subsequent recall of the same object or scene (Altmann 2004; Holšánová et al. 1999; Humphrey and Underwood 2008; Laeng and Teodorescu 2002). Results showing that mental visualizing can be altered or impaired if oculomotor behavior restrictions are applied during scene recollection, or if the eyes move in an image-irrelevant way, have provided additional backing to this theory, for they suggest that scanpaths might provide a motor-based coordinate system to support image generation (Laeng et al. 2014; Johansson and Johansson 2014; Johansson et al. 2012). Additional work has shown that eye movements during mental imagery can be executed independently of how they are produced during encoding, and has claimed that it would be too strong an assumption to conclude that they are reenactments of an original perceptual phase (Johansson et al. 2006, 2010, Johansson et al. 2011, Johansson et al. 2012). Nevertheless, an appeal to motor behavior may be necessary to support memory load during difficult imagery tasks. In sum, irrespective of whether oculomotor activity occurring during imagery is a reenactment of an original perceptual behavior, findings suggest that eye movements may act as a scaffolding structure to generate mental images.

2.2.2 Simulation Theory

Another approach that explicitly invokes a form of action is simulation theory (Barsalou 1999, 2008; Hesslow 2002, 2012). This theory assumes that cognition is body-based in a non-trivial sense. One way to capture this dependence is the claim that all instances of mental activity (including imagery) are perceptual in their genesis and character, and simulate sensorimotor processes through the reactivation of the same neural regions that are involved in actual perception. To imagine something that occurred in a specific location would make the eyes move, for example, because during a neural simulation the imaginer "reconstructs", even though partially and with some distortions, what was originally perceived. Whatever the representation format images rely upon, imaging something is essentially the same as seeing it, and common activation of sensorimotor functions in the brain's modality specific systems would show that imagery, just like perception, is mediated by patterns of embodied responses. The simulation theory charts a picture of imagery that significantly departs from classic,

amodal approaches (e.g., Pylyshyn), and offers a more parsimonious description than the scanpath theory, not only because reenactments of experience in modality-specific states need not be a complete and accurate reinstatement of the original perceptual one, but because patterns of embodied simulation are taken to play a ubiquitous role, and support all instances of cognitive activity (Barsalou 2008; Barsalou et al. 2003).

2.2.3 Emulation Theory

A final approach involving a form of action is the emulation theory (Grush 1998, 2004). According to this view, the employment of an emulator, viz. a neural circuit internally representing the body or the environment, is pivotal to the accomplishment of certain imagery tasks. Imagine the action of grabbing a corkscrew in order to open a bottle of wine. You can form such an image, the theory states, because an emulator of the body, upon receipt of an efferent copy of the motor command, generates a mock version of the proprioceptive information that the body would produce were the motor command to be effectively executed. Running an internal model of the body (or environment) which can be interacted with, is what the theory takes to ground the capacity to mimic the input–output operations that the real body enacts. A mere motor plan would not suffice to mentally imagine the action of reaching out for the corkscrew to open the bottle of wine. This because actions also involve a sequence of proprioceptive states both in terms of sensations (what it will feel like) and kinematics (where and when the hand will arrive), for which an emulator of the body is required. In sum, in order to account for some key features of imagery, this theory takes neurally implemented models of an agent's body and environment to be necessary ingredients.

2.3 Approaches Involving Action and No Representation

We now come to embodied approaches where the notion of representation is clearly rejected, and the importance of action is emphasized. Within this cluster two distinct views have started to emerge. These are the enactive and the sensorimotor approaches. Despite important commonalities, such as the rejection of passive, disembodied, brain-based notions of perception (and cognition), the two approaches differ significantly with respect to the particular way they address the involvement of action in conscious experience. Since, we think, reasons for their divergent views of imagery stem from their respective account of the role of embodied action in perception, we discuss what notion of embodiment and activity figures within their philosophical position as concerns normal perception before illustrating how their different views motivate distinct accounts of imagery.

3 Action-Bound Approaches to Perception: A Comparison Between Enactive and Sensorimotor Accounts

There are valid reasons for thinking that enactivism and sensorimotor theory should not be lumped together and used interchangeably. Perhaps the most important one is the fact that proponents of the sensorimotor theory (Noë 2010,

Noë 2004[4]; O'Regan and Noë 2001; O'Regan 2011) clearly state that they do not entirely adhere to the enactivist paradigm as it was introduced by Varela et al. (1991). From their point of view, what makes perceptual experience possible is not active engagement with the environment per se, but knowledge of what action brings about when bodily engagements become actualized (i.e., knowledge of sensorimotor contingencies).

A second reason to resist lumping these two approaches together is the fact that the literature itself admits different interpretations of the view that perception involves action. To understand this point and the degree of controversy it has attracted, it will be useful to refer to O'Regan and Noë's BBS article (2001), which contains the official philosophical statement of the sensorimotor theory. Here, the two authors claim that perceptual experience is a mode of exploration of the environment that is mediated by the mastery and exercise of sensorimotor knowledge. This statement about the role of action in perception can be read in two ways, one moderate and one radical, and their tension already began to emerge in the commentaries of the BBS article (2001) and the debate following it (e.g., Loughlin 2014; Shapiro 2010; Hickerson 2007).

On the more radical interpretation, visual experience equates to skillful activity and depends on having actually to move, viz. on performing the actions that reveal which lawful relations are being obeyed when the perceiver moves relative to a visual target. To tie perception to action is to say that what an organism senses is a function of how it moves and that the sensible world can show up in consciousness, viz. be present, revealed, brought forth only through active engagement with it. Though one might reasonably question whether anyone has ever clearly, overtly and consistently committed to such an extreme position, we find that this way of characterizing the role of action in perception echoes to some extent the enactive program (Hutto and Myin 2012; Maturana and Varela 1992; Thompson and Varela 2001; Thompson 2007; Varela et al. 1991). Hutto and Myin (2012) for instance, endorse a strong reading of the embodiment thesis, which equates basic cognition with concrete patterns of dynamic interactions between organisms and the world, and claim that "it is not knowledge (embodied know-how) that gives perceptual experiences their intentionality and phenomenal character; rather, it is the concrete ways in which organisms actively engage with their environments" (2012: 30). They further deflate the mediating role of practical knowledge by adding that in engaging with the environment there is nothing organisms know or need to know.

On a moderate interpretation of O'Regan and Noë's statement, which we take to more accurately reflect the core commitments of the sensorimotor approach, perceptual experience requires that one appreciate the relevance of movement for sensory stimulation, even if no movement is undertaken (Noë 2010; O'Regan 2011). The essential idea thus is that perception depends on *knowledge of what movements bring about* (Noë 2010; O'Regan and Noë 2001; O'Regan 2011).

Emphasis on practical knowledge of possibilities for action, rather than actualization of movement through skillful behavior, makes the sensorimotor approach particularly

---

[4] Alva Noë in his 2004 single-authored book adopts an « enactive » approach and specifies that perception is enactive in that it is constituted by sensorimotor knowledge, namely, practical grasp of the way sensory stimulation changes as a function of movement. What he calls « enactive » approach here is no more and no less than the sensorimotor contingency theory. He borrows the term from Varela et al. (1991), but rightly acknowledges that he may not use it in exactly the same way (see p. 233).

suited to account for phenomena that are not easily explained by the standard view of perception as a process consisting in the production of internal representations. One such phenomenon is "virtual presence", the fact that one can be phenomenally conscious of something even though only portions of it are actually in view.

Take a cup and place it on a table in front of you. Step back, stand still and look at it. Although what you see in the restricted sense is only the side facing your eyes, not the back of it, which remains occluded, you visually experience the cup as a whole: you experience a "presence in absence" (Noë 2004, 2006, 2012). Rather than resorting to internally stored representations whose activation generates visual experience, the sensory motor theory appeals to one's implicit grasp of practical conditionals shaping the interplay between the perceiver and the world. We are making the point here that mastery and exercise of sensorimotor knowledge involve entertaining the possibility to be looking at the cup from different vantage points. In this sense, to say that the back of the cup is phenomenally present in perception (not present as imaged) means that we know we would see the part of it that is now occluded from view if we moved in a certain way. The idea is similar to the feel of being at home, the distinctive feel of having everything in one's household easily within reach (O'Regan 2011). To feel at home is to know that there is a variety of actions that can be undertaken, but none of them needs to be really carried out in order for the feel to arise. All that is required is to be *poised* to enact a range of motor capabilities. In the same way, the experience of seeing the cup as a whole, despite our currently only having information about its front side, consists in the *potential* to make contact with the unseen parts. Interestingly, this point finds support in the phenomenon of "boundary extension"[5] (Intraub and Richardson 1989).

There is a point here that is worth clarifying. What sensorimotor knowledge enables one to do is *not* to predict, given a set of available internal hypotheses about the outer world, what things that are out of view are going to be like, but to entertain the possibility to *access* them. The sensorimotor understanding that figures in this type of expectation specifies that it is part of the perceptual experience of a cup that a cup has a back side, leaving open and indeterminate what it should look like. Hence, the claim that visual experience involves one's having sensorimotor expectations does not mean that seeing is (implicitly or explicitly) predicting visual outcomes on the basis of the best set of available internal hypotheses about the world. Rather, it means that one knows that if one were to move around the cup, a back side of one kind or another would come into view. To *expect* thus means to hold faith that visual experience will have a (plausible) sensorimotor profile. Interestingly, Ryle's reference to thimble-seeing contains a similar remark. Someone who looks at a thimble while having a visual sensation, Ryle notes, must have previously learned and not forgotten what a thimble looks like. Knowing how a thimble looks like "he is ready to anticipate, though he need not actually anticipate, how it will look, if he approaches it, or moves away from it; and when, without having executed any such anticipations, he does approach it, or move

---

[5] Boundary extension is an instance of false memory, whereby people report having seen parts of a scene that fall outside a picture's boundaries. For example, when asked to draw from memory a previously displayed picture, people report not only what was actually contained in it, but tend to include details as if the picture had been taken from farther away, thus including more parts of a given object. These 'errors' in boundary judgment actually display a very good prediction of how the world really is, and the rapidity and spontaneity with which they occur suggest that sensory information processing embodies spatial expectations resulting from viewpoint shifts.

away from it, it looks as he was prepared for it to look" (1949/Ryle 2009: 208). The upshot is that the perception of a thimble does not involve pondering, thinking over, making conjectures, memory images or mental replicas of thimbles, but rather "having a sensation in a thimble-seeing frame of mind" (1949/Ryle 2009: 209). To be in a thimble frame of mind is to be prospectively prepared for a variety of sights "of none of which need the thought actually occur to him" (1949/Ryle 2009: 209).

Summing up, the sensorimotor approach offers a form of enactivism in that it rejects the appeal to an internal model of the world and stresses patterns of interaction. It departs, however, from enactivism in that it suggests that being poised for action is enough. It is in this way of being poised, that perception is grounded in mastery and exercise of embodied know-how.

4 Enactivism vs Sensorimotor Approach to Imagery

Having considered what notion of embodiment and activity figures in the enactive and sensorimotor approaches to normal perception, we move to discuss and compare their respective takes on imagery. We aim to show that we are closer to understanding what imagery is when we say that it involves a particular type of know-how, than when we say that it involves a reenactment of the exploratory activity that takes place in perception (i.e., reenactment thesis).

4.1 The Enactive Approach to Imagery

Inspired by theories developed by Hebb (1968) and Neisser (1967), the enactive proposal is that mental visualization of an object or scene consists in rehearsing or re-creating, at least in part, the perceptual exploratory acts that would be performed if one were actually perceiving whatever is being imaged (Thomas 2014, 2010, 2009, 1999; Thompson 2007). Call this the reenactment thesis.

This thesis builds upon the assumption that action and perception are tightly coupled. This coupling is particularly emphasized by Thomas (2009, 2003, 1999). On his view, just as seeing consists in a schema-guided interrogation of the outside world, so too does mental visualizing, which reenacts, in an off-line fashion, the activity of exploring and probing the environment that is typical of perception. Imagery is thus experienced when the same schema directing and instructing the perceptual apparatus during actual seeing is allowed partial control. As Thomas (1999: 223) puts it, "during imagery the schema is active in much the same way that it is during perception. It still sends out at least some of its 'orders' to the perceptual instruments, and selects procedural branches to follow".

According to this view, therefore, when we imagine something, we do not experience a mental picture, and there is no internal mental representation or construct towards which the rehearsed perceptual exploratory behavior is directed. There is only the *activity* of imaging, which reenacts the stimulus-dependent patterns of motor behavior that would take place if the imaged thing were actually being looked at. Borrowing Thomas' example (2014), we have imagery of a cat "when we go through (some of) the motions of looking at something and determining that it is a cat, even though there is no cat (and perhaps nothing relevant at all) there to be seen."

Research on oculomotor behavior has provided reasonable backing for this thesis, and evidence of spontaneous eye movements that closely reflect the content and spatial features of a previously observed object or scene has been taken to confirm that sensorimotor reenactments play a central role in reasoning about absent, distal stimuli (Johansson and Johansson 2014; Johansson et al. 2005, 2006, 2010, 2012; Laeng et al. 2014; Fourtassi et al. 2013; Spivey and Geng 2001; Spivey et al. 2000, to name but a few). Despite some apparent similarities with the scanpath theory, there are important distinctions from a theoretical standpoint. Thomas' view, prototypical of the enactive approach, is anti-representationalist, and does not consider findings about oculomotor activity as evidence that eye movements operate as spatial markers to construct an inner, coherent picture of whatever is being imaged.

Summing up, the idea behind the enactive account is that imagery is an offline, environmentally-decoupled active search for information, and requires mentally re-hearsing the ways in which we visually explore and probe the environment. It is important to note that "action" rehearsal plays a pivotal role in this approach: what imagery requires is not assembling information into an appropriate representational format, but *active interrogation* of the environment supported by the same "schemata" that during actual perception specify how to direct attention to an object or scene (Thomas 1999, 2014).

4.2 Challenges to the Enactive Approach to Imagery

There are potential worries and ambiguities in modelling imagery in motoric, non-representational terms. One general objection concerns the very possibility to do without representations when referring to something that is absent, distal or inexistent. Clark (1997) (see also Clark and Toribio 1994), raising this issue, points out that imagery, dreaming, planning and reasoning about counterfactual states of affairs are representation-hungry cognitive tasks, and heavily depend on internal "stand-ins" or surrogates for the absent phenomena. A second, related objection is that if imagery is used to solve problems, like in mental rotation tasks, reenacting a corresponding overt perceptual exploratory behavior is insufficient, for by itself it cannot explain some key features of imagery (Foglia and Grush 2011). To understand why the enactive approach falls short unless it appeals to representations, consider a situation in which you are presented with pairs of drawings featuring three-dimensional geometrical figures. In each pair of drawings the figures are rotated in a slightly different way and you are asked to check whether the one on the right can be brought into correspondence with the one on the left. Independently of the time taken to confirm whether the figure on the right is congruent to the reference one, to solve the task you will have to perform a mental rotation. We take enactivists to agree that imagery is used to solve problems of this kind, and to claim that through a reenactment of a corresponding *overt* perceptual behavior the solution to the task becomes apparent.

What might such an *overt* perceptual exploration be? One possibility would be an actual manual rotation of the geometrical figure on the right until it is clear whether or not it matches with the one on the left. This *overt* perceptual exploratory behavior has two distinct elements: a motor component (manual grasp and hand rotation), and a target which the overt behavior can act upon, namely, the geometrical figure the motor action is directed towards. Crucially, both components are necessary, and both are

present in overt behavior. An empty hand rotation would not make the solution to the problem perceptually apparent.

Now, what happens when we try to solve the problem using imagery? The enactive proposal is that we covertly engage the same active behavior that would be carried out in the overt condition. This means that we solve the mental rotation task by reproducing, in an off-line fashion, the same motions that would be produced if acting overtly. This seems insufficient, just as an empty hand rotation is insufficient to make the solution perceptually apparent in the overt case. What is needed in this imagery condition is something corresponding to the real geometrical figure, that is, an internal model (i.e., representation) which a set of active behaviors can be directed towards. The upshot is that, if we have correctly interpreted the reenactment thesis as involving purely reenacting an observer's exploratory behaviour itself, then, enactivists would have to accept a representationalist view that has traditionally been the focus of their opponents. A way out would be to give up this "motocentric" approach and abandon any hope of finding a satisfactory theoretical alternative to representationalism in pursuing the enactive proposal.

Given these criticisms of the reenactment thesis, an alternative way to make progress in understanding imagery may be to adopt a different stance and reject the claim that mental visualizing consists in going through the motions of an active interrogation and exploration of the environment in the absence of the object (Thomas 1999, 2014). Under the stance that we adopt, perception is not, as implied by the enactive (but also classic) approach, a benchmark, a standard against which imagery is assessed as being of an inferior species, an impoverished form of seeing. Instead, we emphasize their similarities: seeing is almost as poor as imaging[6] and both are essentially the same because they involve exercising the same implicit, acquired practical knowledge concerning the potential applicability of the sensorimotor laws associated with a given sensation. The main difference is that in visual interactions, there are physically realizable properties that are manifest, and these account for the fact that sensations have a real sensory presence, whereas imagery does not. One among these properties is the fact that bodily movements really do cause sensory changes, whereas when one is imaging, bodily movements may have no effect on the content of one's imagery (O'Regan and Noë 2001; O'Regan 2011 call this "bodiliness"). Another is the fact that sudden external events (like a bright flash) really might incontrovertibly grab one's cognitive system ("grabbiness"). Another is the fact that sensory input really can change without voluntary control by the observer, as when someone really does turn the cup we are looking at ("insubordinateness"). Finally, in the case of really seeing, any amount of detail of the object being seen is immediately available by the slightest flick of the eye or of attention. Whereas when one is imaging, detail must be constructed mentally and is not immediately available ("richness").

To epitomize the sensorimotor view, then, one could say that imaging is fundamentally the same process as seeing, and differs from it only to the extent that seeing is additionally enriched by retinal content, by bodiliness, grabbiness, and insubordinateness. There are of course important differences between the two processes, and these consist in differences between real and unreal engagement.

---

[6] The idea that the phenomenology of seeing is much less rich than is usually thought is supported by such phenomena as change blindness (Simons and Levin 1997) and inattentional blindness (Mack and Rock 1998).

4.3 The Sensorimotor Approach to Imagery

In contrast to the enactive approach, we propose to consider imagery to be constituted by the exercise of knowledge of the potential applicability of the laws that account for the dependencies between sensory input and motor output. .

Imaging a cup thus is to be currently making use of the fact that one is *familiar with* and poised to exercise (practical) knowledge of the general patterns of sensorimotor dependencies that are typical of a cup-type interaction, such as the fact that it looms large in the visual field when we physically approach it, and that its hidden parts become accessible through potential exploratory movements. What one imagines, under this circumstance, does not depend on engaging an activity of looking at "nothing" by partly reproducing the perceptuo-motor behavior that would take place were the imaged object to be actually present, but it consists in being in a state of "attunement" to the sensorimotor laws that have been previously established (learned). "Attunement" to a sensorimotor law refers to a state of currently being poised to confirm a previously acquired familiarity with the ways in which sensory input would potentially change as a function of possible movements. The occurrence of such a state of attunement indicates that a law-like relationship is applicable, and such attunement can occur even when actual movement and sensory input are missing. It is worth emphasizing that the type of knowledge underlying this attunement is implicit, practical knowledge, and not cognitive "knowledge that".

In sum, in the sensorimotor approach, imaging involves being mentally poised to rehearse exploration of an object, but without actually rehearsing that exploration. This is in contrast to the enactive account, where effective (mental) exploration is required.

Empirical evidence supporting the view that perception and imagery share the same neural mechanisms (Ganis et al. 2004; Kosslyn et al. 1993, 1997) does not go against the sensorimotor account. The sensorimotor account can explain the common neural processing as an epiphenomenal effect triggered by the possession and exercise of sensorimotor know-how. For instance, imaging a cup would give rise to brain activation in vision-related areas because mastery and exercise of sensorimotor know-how is enough to induce a state similar to the one corresponding to actually perceiving a cup, with the necessary sensorimotor dependency associated with the sensation being currently applicable. To stress this point further: what is being activated during an imagery experience is not a specific perceptual event or an underspecified set of events and conditions, but rather the knowledge of the potential applicability of the law describing the events corresponding to it.

Having argued that sensorimotor laws underpin all forms of imagery and perception, we can thus understand how it might easily come about that a person would have episodes of imagery that feel vivid, present and almost real. The person simply has to rely on his sensorimotor knowledge and implicitly expect that the sensorimotor laws associated with a sensation can apply. If we wish to image greater details, such as whether the cup is red or green, large or small, dented or brand new, we can also further narrow down the laws so that they apply to more and more detailed aspects of what we are imaging. To do this we simply have to implicitly assume (although incorrectly) that if the eyes were to move we would be able to confirm which of the above conditions is satisfied. What is special about this example is that it suggests that imagery can be highly detailed and yet completely indeterminate in the relevant way. It

is important to note that this accessibility to more details is enabled by the same mechanism that, in visual perception, underlies the feel of having everything within view. So just as in visual perception we are under the impression of continuously accessing aspects of the environment by directing our attention, so too in mental imagery are we under the impression of accessing any detailed information whenever we want it. This state of 'imaging any details' depends on being able to invoke at any moment past, established knowledge of the laws governing the correlation between movement and sensory stimulation.

5 Final Remarks

With the shift from traditional to embodied cognitive science, a new imagery debate has emerged.

Within the embodied approach, however, two quite different accounts have too often been considered interchangeable. These are the enactive and the sensorimotor views. In this paper we have argued that, despite some commonalities, these two views have distinct theoretical traits, and these become manifest when the notion of sensorimotor interaction is critically examined. We started by examining the differences between sensorimotor and enactive accounts of normal perception, and then applied them to imagery. We then questioned the plausibility of the reenactment thesis, the claim that perceptual behavior is recreated during imagery. We rejected this thesis because it yields internal inconsistencies, for in order to work, it would have to accept a view (representationalism) that has traditionally been espoused by its opponents.

In line with the sensorimotor approach, we proposed to refer to imagery as consciously accessing the laws that account for the dependencies between sensory input and motor output. The sensorimotor approach thus stresses the explanatory role played by the possession and use of sensorimotor knowledge, and suggests that "attunement" to sensorimotor laws, *not* reenactment of perceptual experience, is essential to imagery. Having argued that sensorimotor knowledge (know-how) under-pins all forms of perception and imagery, we can thus understand how it might easily come about that a person's imagery might feel vivid, present and almost real. Vivid imagery of drinking from a cup, for instance, would consist in holding implicit expectations regarding how sensory stimulation changes as a function of movement, and this type of know-how would suffice to have almost the same experience as one has when knowledge of the sensorimotor laws associated with a cup are currently really being obeyed. To image thus is to rely on implicit, unfulfilled expectations, which however cannot be sustained for long, precisely because no real engagement with the world (one that possesses properties such as bodiliness and grabbiness) actually occurs. From a neurophysiological point of view, this state of "holding implicit expectations" would be much easier to generate than having to reenact somewhere in the brain the stimulus-dependent patterns of motor behavior that would take place if the imaged thing were actually being looked at.

References

Altmann, G.T.M. 2004. Language-mediated eye movements in the absence of a visual world: The "blank screen paradigm". *Cognition* 93(2): B79–B87.

Barsalou, L.W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22(4): 577–660.

Barsalou, L.W. 2008. Grounded cognition. *Annual Review of Psychology* 59(1): 617–645.

Barsalou, L.W., W. Kyle Simmons, A.K. Barbey, and C.D. Wilson. 2003. Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences* 7(2): 84–91.

Brandt, S.A., and L.W. Stark. 1997. Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience* 9(1): 27–38.

Chemero, A. 2009. *Radical embodied cognitive science*. Cambridge: MIT Press.

Clark, A. 1997. *Being there: putting brain, body, and world together again*. Cambridge: MIT Press.

Clark, A. 2008. *Supersizing the mind: embodiment, action, and cognitive extension*. New York: Oxford University Press.

Clark, A., and J. Toribio. 1994. Doing without representing. *Synthese* 101(3): 401–431.

Foglia, L., and R. Grush. 2011. The limitations of a purely enactive (Non-representational) account of imagery. *Journal of Consciousness Studies* 18(5–6): 35–43.

Fourtassi, M., Hajjioui, A., Urquizar, C., Rossetti, Y., Rode, G., and Pisella, L. 2013. Iterative fragmentation of cognitive maps in a visual imagery task. (F. P. de Lange, Ed.) *PLoS ONE* 8(7).

Gallagher, S. 2005. *How the body shapes the mind*. Oxford: Clarendon.

Ganis, G., W.L. Thompson, and S.M. Kosslyn. 2004. Brain areas underlying visual mental imagery and visual perception: An fMRI study. Brain Research. *Cognitive Brain Research* 20(2): 226–241.

Grush, R. 1998. Perception, imagery, and the sensorimotor loop. In *A consciousness reader,* Eds. Esken, F. and Heckman F.-D. Schoeningh Verlag.

Grush, R. 2004. The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences* 27(3): 377–396.

Hebb, D.O. 1968. Concerning imagery. *Psychological Review* 75(6): 466–477.

Hesslow, G. 2002. Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences* 6(6): 242–247.

Hesslow, G. 2012. The current status of the simulation theory of cognition. *Brain Research* 1428: 71–79.

Hickerson, R. 2007. Perception as knowing how to act: Alva Noë's action in perception. *Philosophical Psychology* 20(4): 505–517.

Holšánová, J., B. Hedberg, and N. Nilsson. 1999. Visual and verbal focus patterns when describing pictures. In *Current Oculomotor research*, ed. W. Becker, H. Deubel, and T. Mergner, 303–304. New York: Springer US.

Humphrey, K., and G. Underwood. 2008. Fixation sequences in imagery and in recognition during the processing of pictures of real-world scenes. *Journal of Eye Movement Research* 2(3): 1–15.

Hutto, D.D., and E. Myin. 2012. *Radicalizing enactivism: Basic minds without content*. Cambridge: MIT Press.

Intraub, H., and Richardson, M. 1989. Wide-angle memories of close-up scenes. J*ournal of Experimental Psychology: Learning, Memory and Cognition* 15: 179–187.

Johansson, R., and M. Johansson. 2014. Look here, eye movements play a functional role in memory retrieval. *Psychological Science* 25(1): 236–242.

Johansson, R., Holšánová, J., and Holmqvist, K. 2005. What do eye movements reveal about mental imagery? Evidence from visual and verbal elicitations. Proceedings of the 27th Cognitive Science conference.

Johansson, R., J. Holšánová, and K. Holmqvist. 2006. Pictures and spoken descriptions elicit similar Eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science* 30(6): 1053–1079.

Johansson, R., Holšánová, J., and Holmqvist, K. 2010. Eye movements during mental imagery are not reenactments of perception. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society,* Eds. Ohlsson S. and Catrambone R. Cognitive Science Society.

Johansson, R., Holšánová, J., and Holmqvist, K. 2011. The dispersion of eye movements during visual imagery is related to individual differences in spatial imagery ability. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, 1200–1205. Austin: Cognitive Science Society.

Johansson, R., J. Holšánová, R. Dewhurst, and K. Holmqvist. 2012. Eye movements during scene recollection have a functional role, but they are not reinstatements of those produced during encoding. *Journal of Experimental Psychology: Human Perception and Performance* 38(5): 1289–1314.

Kind, A. 2001. Putting the image back in imagination. *Philosophy and Phenomenological Research* 62(1): 85–109.

Kosslyn, S.M. 1980. *Image and mind*. Cambridge: Harvard University Press.

Kosslyn, S. M. 1994. *Image and brain*. Cambridge: MIT Press.

Kosslyn, S.M. 2005. Mental images and the brain. *Cognitive Neuropsychology* 22(3): 333–347.

Kosslyn, S.M., N.M. Alpert, W.L. Thompson, V. Maljkovic, S.B. Weise, C.F. Chabris, S.E. Hamilton, et al. 1993. Visual mental imagery activates topographically organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience* 5(3): 263–287.

Kosslyn, S.M., W.L. Thompson, I.J. Kim, and N.M. Alpert. 1995. Topographical representations of mental images in primary visual cortex. *Nature* 378(6556): 496–498.

Kosslyn, S.M., W.L. Thompson, and N.M. Alpert. 1997. Neural systems shared by visual imagery and visual perception: a positron emission tomography study. *NeuroImage* 6(4): 320–334.

Kosslyn, S.M., A. Pascual-Leone, O. Felician, S. Camposano, J.P. Keenan, W.L. Thompson, G. Ganis, et al. 1999. The role of area 17 in visual imagery: Convergent evidence from PET and rTMS. *Science (New York, N.Y.)* 284(5411): 167–170.

Kosslyn, S.M., W.L. Thompson, and G. Ganis. 2006. *The case for mental imagery*, vol. vi. New York: Oxford University Press.

Laeng, B., and D. Teodorescu. 2002. Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science* 26(2): 207–231.

Laeng, B., I.M. Bloem, S. D'Ascenzo, and L. Tommasi. 2014. Scrutinizing visual images: The role of gaze in mental imagery and memory. *Cognition* 131(2): 263–283.

Loughlin, V. 2014. Sensorimotor knowledge and the radical alternative. In *Contemporary Sensorimotor Theory, Studies in Applied Philosophy, Epistemology and Rational Ethics,* Eds. Bishop J.M. and Martin A.O., 105–116. Springer International Publishing.

Mack, A., and I. Rock. 1998. *Inattentional blindness*. Cambridge: MIT Press.

Mast, F.W., and S.M. Kosslyn. 2002. Eye movements during visual mental imagery. *Trends in Cognitive Sciences* 6(7): 271–272.

Maturana, H.R., and Varela, F.J. 1992. *The tree of knowledge: the biological roots of human understanding*. Cognition.

Moulton, S.T., and S.M. Kosslyn. 2009. Imagining predictions: mental imagery as mental emulation. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* 364(1521): 1273–1280.

Neisser, U. 1967. *Cognitive psychology*. East Norwalk: Appleton.

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

Noë, A. 2006. Experience without the head. In *Perceptual experience*, ed. T. Gendler and J. Hawthorne, 411–433. Oxford: Oxford University Press.

Noë, A. 2010. Vision Without Representation. In N. Gangopadhyay, M. Madary, & F. Spicer (Eds.), *Perception, Action, and Consciousness: Sensorimotor Dynamics and Two Visual Systems*, 245–256. Oup Oxford.

Noë, A. 2012. *Varieties of presence*. Cambridge: Harvard University Press.

Noton, D., and L. Stark. 1971a. Scanpaths in eye movements during pattern perception. *Science (New York, N.Y.)* 171(3968): 308–311.

Noton, D., and L. Stark. 1971b. Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research* 11(9): 929–942.

O'Regan, J.K. 2011. *Why red doesn't sound like a bell: understanding the feel of consciousness*, 1st ed. New York: Oxford University Press.

O'Regan, J. K., and N. Block. 2012. Discussion of J. Kevin O'Regan's "Why Red Doesn't Sound Like a Bell: Understanding the Feel of Consciousness." *Review of Philosophy and Psychology*, 3(1): 89–108.

O'Regan, J.K., and A. Noë. 2001. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences* 24(05): 939–973.

Pylyshyn, Z.W. 1981. The imagery debate: Analog media vs. Tacit knowledge. *Psychological Review* 88: 16–45.

Pylyshyn, Z.W. 2002. Mental imagery: In search of a theory. *Behavioral and Brain Sciences* 25(2): 157–182.

Pylyshyn, Z.W. 2003. *Seeing and visualizing: It's not what you think*. Cambridge: A Bradford Book.

Ryle, G. 2009. *The concept of mind: 60th anniversary edition*. London: Taylor & Francis.

Shapiro, L. 2010. *Embodied cognition*. New York: Routledge.

Simons, D.J., and D.T. Levin. 1997. Change blindness. *Trends in Cognitive Sciences* 1(1): 241–282.

Spivey, M.J., and J.J. Geng. 2001. Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological Research* 65(4): 235–241.

Spivey, M.J., M.J. Tyler, D.C. Richardson, and E.E. Young. 2000. *Eye movements during comprehension of spoken scene descriptions. Proceedings of the 22nd annual conference of the cognitive science society*, 487–492. Mahwah: Lawrence Erlbaum Associates Inc.

Thomas, N.J.T. 1999. Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science* 23(2): 207–245.

Thomas, N.J.T. 2003. New support for the perceptual activity theory of mental imagery. Resource document : http://www.imagery-imagination.com/newsupa.htm

Thomas, N.J.T. 2009. Visual imagery and consciousness. In *Encyclopedia of Consciousness,* Ed. W. P. Banks.

Thomas, N.J.T. 2010. Imagery and the coherence of imagination. *Journal of Philosophical Research* 22: 95–127.

Thomas, N. J. T. 2014). Mental imagery. In *The Stanford Encyclopedia of Philosophy,* Ed. E.N. Zalta. Resource document : http://plato.stanford.edu/archives/fall2014/entries/mental-imagery/

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

Thompson, E., and F.J. Varela. 2001. Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Sciences* 5(10): 418–425.

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

White, A.R. 1990. *The language of imagination*. Oxford: Blackwell.