



Coupling to Variant Information: an Ecological Account of Comparative Mental Imagery Generation

Matthew Sims¹ 

Published online: 20 December 2019
© The Author(s) 2019

Abstract

Action-based theories of cognition place primary emphasis upon the role that agent-environment coupling plays in the emergence of psychological states. Prima facie, mental imagery seems to present a problem for some of these theories because it is understood to be stimulus-absent and thus thought to be decoupled from the environment. However, mental imagery is much more multifaceted than this “naïve” view suggests. Focusing on a particular kind of imagery, comparative mental imagery generation, this paper demonstrates that although such imagery is stimulus-absent, it is also stimulus-sensitive. Exhibiting stimulus-sensitivity is sufficient for a process to qualify as coupled to the environment. The notion of *variant coupling* is explicated as the coupling of a cognizer’s perceptual system to variant environmental information. By demarcating the categories of stimulus-absent and stimulus-sensitive cognition, and variant and invariant coupling, this paper expands the conceptual apparatus of action-based theories, suggesting not only a way to address the problem that comparative mental imagery generation presents, but perhaps a way to account for other forms of imagery too.

1 Introduction

Action-based theories of cognition (Gibson 1979; Varela et al. 1991; Clark 1997/2016; Ballard et al. 1997; Cisek and Kalaska 2010; Pezzulo and Cisek 2016) conceive of perception, cognition, and action as a continuous and mutually influencing process, the ‘function’ of which is to guide interaction with an everchanging yet structured environment. Action-based theories may be contrasted to traditional serial information processing views, which conceive cognition primarily as a means of constructing an accurate description of environmental states. Although there is disagreement amongst action-based theories as to how (and the degree to which) agent-environment

✉ Matthew Sims
s1670536@ed.ac.uk

¹ School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Dugald Stewart Building, Charles Street 3, Edinburgh EH8 9AB, UK

interaction is involved in the emergence of cognitive states, these approaches are unified in the importance they bestow upon coupled interaction with the environment. Some forms of action-based theory (e.g., traditional ecological approaches and some enactive theories), taking a wide explanatory scope, attempt to account for the majority of cognitive processes in terms of *ongoing* agent-environment coupling.¹ At first blush, these wide-scope action-based theories are faced with a serious problem when attempting to account for mental imagery. Like other offline phenomena, mental imagery provides cognizers with the means to explore possible behavioural outcomes prior to engaging in action, thus arriving at an ‘optimal’ behavioural choice (relative to one’s goals) in an energy efficient manner. Given its status as an offline phenomenon, mental imagery is commonly seen as stimulus-absent² and thus decoupled from the environment (Pezzulo 2017; Foglia and Grush 2011). For this reason, any action-based account of imagery in which ongoing coupling is deemed essential *seems* deeply flawed. However, mental imagery is much more multifaceted than this “naïve” view suggests.

The multifaceted nature of mental imagery is exemplified particularly well when considering one intriguing form which involves the deployment of imagery in the service of visual comparison tasks with features of perceived objects. For example, one might be given the task to look at an object, generate an image with the same dimensions and then compare image dimensions with those of another perceived object. Such a task, for instance, might involve comparing the size of one box located in the left corner of a room to another similarly sized box located in the right corner of a room in order to know whether or not the leftmost box would fit inside of the other. This particular kind of image generation, what will be referred to as *comparative mental imagery generation* (CMIG), is striking for the reason that, upon close analysis, it challenges the orthodox view in cognitive psychology that all mental imagery generation is environmentally decoupled. In the first part of this paper I will argue that, because the imagery generated during CMIG remains *sensitive to the stimulus dynamics of the environment*, CMIG is both an offline and coupled process.

To do this, this paper argues that CMIG is both influenced by incoming sensory stimuli and may prompt image maintaining behaviour. Possessing this behaviour-eliciting sensitivity to environmental stimuli is sufficient for this kind of mental imagery generation to qualify as what I will call the process of *variant coupling*. Building on the notion of stimulus sensitivity, this paper offers an ecological account of the kind of coupling involved in CMIG. By illustrating how perceptual systems may couple to variant information in the environment, thus allowing for continued imagery maintenance despite ongoing encounters with disruptive stimuli, the ecological account of CMIG demonstrates that action-based accounts should adopt a richer taxonomy of cognitive processes than a standard online/offline bifurcation, if they are to make sense of phenomena like CMIG. By demarcating the categories of stimulus-absent and

¹ Those theories which I am calling wide-scope action-based theories go beyond making the weak claim that agent-environment coupling *mediates* higher forms of cognition (i.e., that higher forms of cognition such as thought or imagination come about in virtue of the covert, ‘as if’ deployment of the same sensorimotor mechanisms that would occur were actual agent-environment coupling were occurring). In contrast, wide-scope action-based theories claim that, like perception and action, many cases of higher cognition emerge from agent-environment coupling.

² More will be said about the notion of stimulus absence below.

stimulus-sensitive cognition, and variant and invariant coupling, this paper expands the conceptual apparatus of action-based theories, suggesting not only a way to address the problem that stimulus-absence presents for comparative mental imagery generation, but also a way that action-based theories may *possibly* account for other forms of stimulus-absent imagery.

This paper is organised as follows: after defining on and offline cognition from the theoretical perspective of action-based cognition, SECTION 2 provides an example and an analysis of CMIG. Using this CMIG example, SECTION 3 introduces the notion of *stimulus-sensitivity* and argues that exhibiting stimulus-sensitivity is sufficient for a mental state/process to be coupled to the environment. SECTION 4 proposes an ecological account of *variant coupling* involved in CMIG. SECTION 5 sketches a more fine-grained taxonomy of offline and online cognitive phenomena based upon the notions of variant and invariant coupling/decoupling and summarises the theoretical gains that this taxonomy affords for action-based theories.

2 Online and Offline Cognition

Within the context of action-based theories, as in the rest of cognitive psychology, the terms *online* and *offline* are readily accepted to demarcate two distinct cognitive modes (Clark 1997; Pezzulo 2017; Bickerton 1996). Online cognition is characteristically understood to be *stimulus-based* (ibid.) or equally, causally dependent upon the task relevant stimulus features in the environment. One might understand causal dependence here in the context of experiential states or processes as follows:

Some psychological state, S , is causally dependent on state of affairs, O if:

- a. Were O not present, S would have failed to arise &
- b. were there a registerable difference in O , then there would be a difference in S

Where (a) expresses a Lewisian (Lewis 1973) counterfactual condition on causal dependence and (b) narrows the relevant causal dependence down to covariation of psychological states with environmental states by introducing a counterfactual conditional that takes a registerable difference³ in O as an antecedent and a difference psychological state as a consequent.⁴

Importantly, action-based theories characterize online cognition as a kind of causally coupled interaction with the environment. An agent may be said to be coupled to the environment when her behaviour with respect to some task-relevant environmental feature brings about changes in the environment (or changes in relation between environmental features and the agent), which in turn modulate that agent's sensory states (i.e., input) acting to guide subsequent behaviour and constraining subsequent

³ Registerable difference is determined by the sensitivity of the type of perceptual system in question. It will suffice for the purposes of this exposition to assume this to be individuated by species.

⁴ This is a mere suggestion as to how one may construe causal dependence of experiential states upon environmental states. If one is not satisfied with these particular conditions, then one is free to replace them with one's best conditions on causal dependence.

sensory states.⁵ Coupling requires mutually ongoing casual activity between two or more systems over time. In the case of online cognition, those two systems are the external environment and the cognizer whose continuous state changes—understood respectively as *environmental states* and *internal, action and sensory states*—are mutually influencing and hence provide mutual information about one another (Jost 2015).

Although there is disagreement amongst action-based theorists as to the details of how offline phenomena come about, it is agreed upon that offline cognition refers to phenomena which are *stimulus-absent*. The generation of the mental activity yielding offline phenomena is said to be largely spontaneous and thus both *causally independent* and *decoupled* from task relevant stimulus features in the environment. It follows that in such episodes of cognition offline states and environment states fail to provide mutual information about one another given that the activity resulting in offline states stops short of acting on the environment and receiving environmental feedback. Whilst offline states such as dreaming and contemplating non-existent entities (e.g., having thoughts about unicorns) might fail—at least in any direct way—to hinge upon prior environmental challenges encountered, many offline phenomena may be seen as “systematic explorations to problems set by experience” (Gerrans 2007, p. 46). In other words, it is typical of a wide range of offline phenomena that they are deployed in solving problems which a cognizer may initially be presented with online. They provide an efficient way to explore behavioural options—often future oriented—thus avoiding costs in the form of time and energy-consuming, ‘real-time’ action or environmental exploration. Some further examples of offline phenomena are:

- *Mentally navigating* a route which has previously not been taken.
- *Mentally rehearsing* a sequence of dance steps without carrying them out.
- *Imagining* how a particular table would look in one’s parlour prior to buying it.

With the online/offline distinction to hand, let’s move onto CMIG. CMIG is an example of an interesting form of mental imagery which illustrates that not all imagery generation is decoupled from the environment. In closely examining CMIG, it will be suggested that the stimulus-absent vs stimulus-based means of analysing cognition overlooks an important concept, that of *stimulus-sensitivity*. Because mental imagery generation may be consistently stimulus-sensitive and stimulus-absent, and (I will argue) stimulus-sensitivity is a way of being causally coupled to the environment, comparative mental imagery generation simultaneously involves decoupling and coupling. This is a significant result particularly for wide-scope action-based theories given that it provides a means of substantiating the claim that higher cognitive states like mental imagery involve *ongoing* agent-environment coupling.

2.1 Comparative Mental Imagery Generation

To understand what CMIG is, imagine a disk pairing game. You are presented with a table upon which are two rows of five disks of various sizes. The two rows are

⁵ Mathematically, this may be described in terms of a series of differential equations in which the changes in any one of the variables that is a constituent of one equation are reciprocally describable as a function of the changes in the other equation’s variables. See Chemero (2004) for an in-depth account of agent-environment coupling using the variable *tau*.

separated by a distance of one metre (or any approximate distance where you cannot view the rows simultaneously by saccades alone). Some of the disks are of noticeably different sizes and others with a mere difference of less than a centimetre (see Fig. 1). The aim is to form pairs of disks from each row by sameness of size as quickly as possible. If you were to compare the size of two discs by actually placing them adjacent to (or atop) one another and looking for a noticeable difference, you would be engaging in real-time, online cognition. Offline cognition, however, might take the form of your *mentally comparing* the size of one disc with another by holding an image of the disk in mind, while looking at the disk dimensions in the adjacent row.⁶

Let's focus on some of the characteristics of offline cognition that are brought out in this example. Introspection reveals the following: during the game, a disk's dimensions (i.e., size) are perceptually sampled, an image is generated and held fixed, transposed to a different location and tacitly compared to the size of a disk that is perceived online. Thus, despite their being derived from a perceptual encounter with the disks, the disk images elicited are *stimulus-absent*. That is, once the disk is perceptually sampled, the maintained image becomes decoupled from the disk it was sampled from, and thus no longer under the causal influence of the disk. For instance, subsequent to sampling and maintenance, one may certainly break the disk without breaking the image. Importantly, this is illustrative of a case in which the mechanisms for generating and maintaining stimulus-absent images are deployed simultaneously to those that are being used for online perception; a capacity that falls out of the fact that mental imagery in a particular modality and perception in the corresponding modality may occur concurrently (Nanay 2017). For the disk image to be used in a comparison task requires both the image and disk percept to be used concurrently.

Such simultaneous deployment should be expected given the general consideration that offline cognition is often used in the service of solving problems which an organism encounters in real-time (Gerrans 2007) and that real-time problems often require constant reference to those environmental features from which the problems stem. Furthermore, it has been established by a range of neuroscientific studies that imagery and perception deploy much of the same neural circuitry (Kosslyn, Kossylin 1978, Kossylin 1980, Kossylin 2005; Farah 2000), supporting what has become known as the continuum thesis.⁷ This thesis, which is assumed true within the cognitive science of mental imagery (Thomas 1999), often begins with the evolutionary assumption that more complex cognitive capacities such as mental imagery generation developed from simpler perception-action capacities (Cisek & Kalaska 2010; Pezzulo and Cisek 2016). At least one version of the thesis then goes on to claim that mental imagery generation makes use of some of the same processing that is deployed in

⁶ For a classic comparative imagery study involving mental rotation, see Shepard and Metzler (1971). Although, it involves 3-D manipulation, something which is much more complex than holding the disk size fixed and transposing it across one spatial dimension for comparison, the Shepard & Metzler mental rotation task may be nonetheless considered a form of CMIG. This is because the mental rotation task, although it involves 3-D rotation, requires holding a generated shape fixed in order to compare it to another shape that is seen online. Rotation may be seen as another manner of mentally transposing an image, albeit a spatially complex manner placing a higher requirement upon working memory.

⁷ The continuum thesis that will be referred to throughout this paper should not be confused with the more controversial *philosophical continuum thesis*. The philosophical continuum thesis states that there is no qualitative difference between mental imagery and perception, but rather a mere quantitative difference. I am greatly indebted to an anonymous reviewer for bringing this distinction to my attention.



Fig. 1 Disk Pairing Game: the aim is to pair the disks from each row according to sameness of size as quickly as possible

perception with the exception that in imagery generation the processing that would normally lead to perception is somehow inhibited (Degenaar and Myin 2014).

If imagery makes use of much of the same neural processing as perception, we might expect that one kind of phenomenon could affect the other. This suggestion is supported by the fact that impairments in one domain can often be accompanied by impairments in the other (Farah 2000). Early research by Bisiach and Luzzatti (1978), for example, discovered that patients suffering from unilateral neglect also tend to neglect the corresponding area of their mental imagery. Additional evidence comes from the noting that making random eye movements disrupts visual imagery (Thomas 1999) and imagining a particular type of object involves making similar saccadic eye movements to those that one would make when perceiving that object (Fourtassi et al. 2018).⁸ In the next sub-section, I will attempt to specify just how the stimuli encountered in perception could affect mental imagery.

3 Stimulus-Sensitivity

The disk pairing game exemplifies CMIG. This domain of offline cognition, I will argue, makes it evident that the online/offline taxonomy as based upon the notions of stimulus-based and stimulus-absent is underdeveloped. CMIG is a temporally extended process in which the spatial comparison of generated and held imagery with spatial environmental features requires the use of both off and online mechanisms. Given that perception requires the processing of environmental stimuli, it may be asked whether the stimuli thus processed have any causal effect upon the accompanying maintained visual imagery?⁹ In other words, although the imagery associated with CMIG is stimulus-absent (i.e., causally decoupled from the sampled source) is it possible that the held imagery is sensitive to impinging visual stimuli encountered during the comparison task? It is by posing this question that the notion of *stimulus sensitivity* reveals itself.

Let us define stimulus sensitivity as follows.

⁸ Fourtassi et al. (2018), by applying a bi-dimensional regression model to eye-tracking data collected during an imagery task (i.e., imagining different towns in France) were able to confirm that saccadic gaze patterns closely reflect spatial dimensions of relative town locations.

⁹ This question is different than that of whether or not the general processes involved in perception, such as random eye movements, affect imagery, something which—as pointed out above—has been empirically supported. This question is particularly about the influence of environmental stimuli upon imagery.

A mental process is stimulus sensitive just in case.

- (1) its individuating features are subject to being affected by incoming stimuli and
- (2) it is likely to elicit adaptive-control responses when such stimuli threaten to degrade the imagery being generated/maintained.

To take an example from the disk pairing game, the mental image is a blue disk of certain size and lasting in its vivacity and dimensions over a certain period of time corresponding to the task at hand. Condition (1) on stimulus sensitivity requires that such an image's individuating features be subject to influence by visual stimuli. Where individuating features are those subjective properties of an image which qualify it as the image token that it is. Let's imagine that after sampling one of the disks on the left-hand side of the table and holding that image, an intense flash of light occurs, hitting your eyes as you direct your head and eyes to the disks on the right-hand side of the table. Were this light flash to affect the blueness, size, or shape of the held image then such an image would satisfy condition (1).¹⁰ Condition (2) requires an image to be likely to elicit behavioural responses that aid in the maintenance of those individuating features of an image. The notion of adaptive-control here is one borrowed from control theory (Pezzulo and Cisek 2016; Ashby 1952). It may be roughly thought of as a means to keep a certain control variable in a system within a restricted range by actively controlling the feedback into the system. In the disk pairing game, the control variable may be construed as one of the image individuating features of any sampled disk. Adaptive control responses would be those which the imager would deploy (tacitly or non-tacitly) in order to maintain that feature during the comparison task. If there were certain behaviours that were likely to be elicited for this purpose, then condition (2) would be satisfied. Are there any reasons to believe that mental imagery satisfies either of these conditions on stimulus sensitivity? Let's look at each condition respectively.

Lending support to the satisfaction of condition (1) we may look to luminance effects studies by Keogh and Pearson (2014). It is well established that when a subject is told to generate a mental image of one of two binocular rivalry images, say a green Gabor patch, and then shown a binocular rivalry display with opposing red and green Gabor patches, the congruent green patch in the display will be experienced as more dominant (Keogh and Pearson 2011; Sherwood and Pearson 2010). By increasing background luminance during the generation of a mental image, Keogh & Pearson have shown that congruent images presented subsequently in a binocular rivalry display were less dominant for good imagers. This suggests that "incoming visual information, which has obligatory access to early visual areas, can interfere with [...] internal sensory representations" (Keogh and Pearson 2014:10). In other words, online stimuli can—in some cases—degrade imagery; an unsurprising result, given the continuum thesis described above (Thomas 1999).

What about condition (2)? One way to frame the question about satisfying (2) is to ask *what can an agent do to avoid image degradation?* In responding to this question, I will focus upon two kinds of responses: saccadic eye movement and pupilar

¹⁰ That this could be the case presupposes that an intense flash of light hitting one's eyes typically blurs online vision and that the affected processes are the same exploited by offline vision. I am indebted to an anonymous reviewer for the suggestion to clarify this point in more detail.

constriction. These responses offer indirect—though, I suggest, compelling—support for the satisfaction of (2) based on empirical studies. In coming to understand how saccades may be used to avoid degradation, it is helpful to keep in mind the seminal research on eye tracking and top-down influence by Alfred Yarbus (1967). In one famous experiment, Yarbus, used a painting depicting a family gathered in a parlour as a visual stimulus. After an initial viewing, before which no particular instructions were given, subjects were assigned various tasks to carry out prior to each subsequent 3 min viewing (e.g., “estimate the material circumstance of the family”, “remember the positions of the people and objects in the room”, “Surmise what the family had been doing before the arrival of ‘unexpected visitor’”). Whilst carrying out these various tasks, the subjects’ saccades were tracked. It was found that the saccadic patterns varied significantly with the kind of task assigned, thus providing evidence that *saccades can be driven to pursue meaningful information given the nature of the task*. Saccades, it turned out, are not just passive responses, but are under the influence of ‘top-down’ control, deeply reflecting the nature of the goals to be achieved.¹¹ With this in mind, I would like to suggest a variation on Yarbus’s findings that will underwrite a response to the question about (2) posed above, viz., *saccades can be driven to avoid disruptive visual stimuli given the nature of the task*.

Recent research by Kilpeläinen and Theeuwes (2016) supports this claim. After establishing a penalty zone (an area which subjects would be financially penalized for fixating upon) in a visual display and instructing subjects to locate cues that would be intermittently flashed in the same display, Kilpeläinen & Theeuwes found that saccades can be flexibly adapted to changing environmental circumstances to improve reward outcome. Given that improving reward outcomes in this experiment is equivalent to decreasing penalization outcomes, these results suggest that the bodily exploration of the environment involved in perception might be adaptively shaped both by what is desired and what is not desired: the result being the avoidance of undesired stimuli. Given that this is the case, an account of CMIG as a stimulus-sensitive process predicts that in mental imagery cases subsequent to exposure to image degrading stimuli (e.g., high luminance surfaces), a subject will reduce the number of saccades to locations containing such stimuli in order to improve the task outcome.

Pupilar dilation provides striking additional support for the satisfaction of condition (2). Studies by Binda et al. (2013) have shown that pupils constrict when subjects are merely looking at photos of the sun. These findings suggest contextual expectations elicit pupilar response and modulation of sensory processing. Returning to the context of mental imagery, given these results one might expect that after learning which stimuli (or changes in stimuli) predict the presence of high luminance conditions, an agent’s pupils might constrict to avoid the degradation of imagery and dilate otherwise (say, for further sampling during CMIG). This goes beyond mere anticipatory pre-emption in that some stimulus is required as a predictor of the upcoming increase in luminance (much as the image of the sun serves as a predictor of vision-degrading

¹¹ Although today this conclusion is not surprising, in 1967 it was. Three decades prior to Yarbus’ research into active vision, work by G.T. Buswell (1935) confirmed that the eye movements are under the influence of goals. However, due to the fact that his experiment featured different kinds of tasks requiring different kinds of cognitive faculties (e.g., remembering vs judgement), Yarbus managed to demonstrate not only that goals influence saccades but that differences in goals are systematically reflected in saccadic activity in a fine-grained manner.

brightness). It is this predictor stimulus that is necessary to drive behaviour in ways that help stay within a certain imagery-friendly luminance range. Thus, such constriction and dilation would have to be a measured and delicately balanced response; one influenced both by top-down and bottom-up processing.

Taking into consideration both luminance effects on imagery and evidence that saccades and pupilar responses may be contextually driven to avoid anticipated image-degrading stimuli (or equally driven to couple to stimuli that is non-degrading to imagery), gives strong, though indirect, support for the hypothesis that CMIG is stimulus sensitive. To be sure, there are other instances of visual mental imagery generation in which an agent is not subject to the influence of visual stimuli. For example, were one to wear an eye mask blocking out all visual stimuli, luminance effects (in addition to all other possible light stimuli effects) would be null; there would be no occasion for visually based adaptive-control responses. In such a case, it might well be that visual mental imagery generation is *stimulus-insensitive*.¹² However, since CMIG by definition involves comparing imagery with perceived objects and thus requires an openness to incoming visual stimuli, the imagery generated and maintained during CMIG remains *sensitive* to visual stimuli which, via adaptive control, are kept within a limited, image-conducive, range.¹³ Stimulus sensitivity thus implies that the processing underlying CMIG remains causally coupled to the environment. CMIG piggybacks on stimulus-based perception but is a distinct offline process.

Crucially, this *opens up the theoretical space* for the images in CMIG to be consistently *both* stimulus-absent (i.e., causally decoupled from the environmental source of their phenomenal character) and yet stimulus-sensitive (i.e., causally coupled to visual stimuli encountered during the visual task). Accepting this kind of analysis of CMIG as plausible, however, is to a large extent dependent upon providing an account of what it is that visual systems are coupling to. In the course of addressing this question in the next section, some important further details about stimulus-sensitivity will emerge.

4 Coupling to Variant Information

What does a cognizer couple to during CMIG? In answering this question, it will be helpful to examine the kind of coupling relation that an ecologically-informed theory of active cognition is committed to. The reason for this is that ecological theories place an emphasis upon the environment as playing an indispensable role in cognition and thus such theories offer the most richly developed framework for thinking about coupling amongst action-based theories. Drawing on Gibson's ecological psychology (Gibson

¹² It is an interesting and open empirical question whether or not (or the degree to which) non-visual exteroceptive modalities such as sound have any significant effect upon visual mental imagery. That this is the case with respect to auditory stimuli seems to me to clearly to be the case, particularly with respect to sound effects. See Miller and Marks (1992) for evidence as to how radio sound effects affect the elicitation of visual imagery.

¹³ This claim is consistent with the further claim that saccades reflect 'encoded' spatial relations and hence are sensitive to various spatial dimensions which have been encoded into memory. The point again is that saccadic behaviour during imagery maintenance is not only driven by top-down, inner-dynamics but also bottom-up, environmental dynamics.

1979) one form of enactivism (Kiverstein and Rietveld 2018; Bruineberg et al. 2017; Rietveld and Kiverstein 2014) understands coupling to be a relation between inner dynamics and the environmental information. If environmental information is given its usual gloss of being invariant patterns that specify the layout of the environment (Gibson 1979), then it is by coupling to invariant information that organisms perceive the environmental layout specified by that information.

With this in mind, there is a clear distinction to be drawn between invariant information and the stimuli involved in stimulus sensitivity. Given the description of stimulus-sensitivity provided above, it seems that the kind of stimuli that affect the phenomenal character of imagery and elicit response does not specify the environmental layout. Returning to the Keogh and Pearson (2014) study, although the high luminance light that disturbs imagery generation originates from a surface, the light itself does not specify that surface. In Gibson's (1979) terminology, such light stimulus is not "ambient" but merely "radiant" and thus falls short of carrying invariant information. Given that this is the case, the latter does not specify unchanging environmental structure. So, whatever the kind of coupling that stimulus-sensitivity implies is, it is distinct from that which occurs in perception.

The coupling that ecologically-informed theories have traditionally used in accounting for perception is understood to involve information that specifies the environmental layout or *what* actions the layout offers (i.e., affordances). However, the environmental *relata* which stimulus-sensitivity involves, because of being behaviour guiding nevertheless serves as information for the perceptual system. In the case of CMIG, the visual system is coupled to meaningful, non-homogenous stimuli that are instructive as to *how to behave* relative to the task at hand.¹⁴ Alone, high luminance radiant light does not serve as information, but given the limited range of available stimulus conditions within which imagery can be maintained *and* the location of the perceiver in relation to varying environmental stimulus conditions, such radiant light stimulus is informative: it specifies where not to saccade from *this* location.

One might object, arguing that such information—call it 'variant information'—seems nonetheless unlikely to be the kind of thing which could be the environmental *relata* that underlies coupling. For it would seem odd that the occasional encounter with high luminance conditions and the elicited response to avoid them could be considered anything other than a temporary coupling and immediate decoupling from such stimulus conditions. In other words, it seems that the visual system is adapting in ways so as to remain *decoupled* from high luminance, image degrading stimuli. In response to this, one need only be reminded that light conditions are continuously fluctuating in ways that are sometimes better or worse for imagery maintenance. During CMIG the visual system is constantly engaged in an attempt to stay coupled to the range of stimulus conditions that are best suited for imagery maintenance and it does this by coupling to and decoupling from degrading stimuli exogenously encountered. Thus, it may be argued, that the visual system does not couple to high luminance surfaces *only*

¹⁴ One way of conceiving the difference between those things that are specified affordances and those things which variant information specify may be brought out respectively in the vocabulary of "action selection" and "action specification" (Pezzulo and Cisek 2016). This may be seen as a difference between selecting amongst those activities (i.e., *what*) the environmental layout offers (e.g., sitting, standing, walking through, etc.) and the specification of an action that describes *how* one's body must behave to successfully bring about a selected action.

to decouple from them, but, instead, when such conditions are encountered, by responding in ways to avoid them (or adapt to them), the visual system remains coupled to the stimulus conditions which are conducive to imagery maintenance. Variant information in this sense guides the visual system back to the ‘optimal’ conditions, constraining the task driven saccadic patterns that unfold during CMIG.

One may think about this process analogously to descending an unfamiliar staircase in the dark by sweeping a foot along until it meets the next stair below. Failing to encounter a surface upon which to shift one’s weight, one continues to sweep and until such a surface is detected. A sweep that fails to detect anything is nonetheless informative as to where not to place one’s weight. On the other hand, when one detects a surface, one knows where to shift one’s weight. Through a series of such moves one descends the darkened staircase, staying balanced upright along the way. Similarly, variant information guides the visual system in both *how* and *how not* to not behave given both the task at hand and the location that one occupies.

The specifying role of this ecological variant information is made explicit in the following passage from Warren (2005), who, drawing a comparison between it and the role of invariant information, writes:

“Reciprocally, the varying spatial relation between objects and perceiver is specified by the *perspective structure* of stimulation. This corresponds to the view of the environment *from here*, which locates environmental surfaces and objects relative to the perceiver, and the perceiver relative to its environment” (Warren 2005, p. 343).

This perspectival structure is what Gibson (1979) referred to as *propriospecific information*. Since it specifies something about the perceiver’s current position relative to the environment, it may be thought of as providing her with the kind of indexical information required to direct her body in ways that *would* change her relationship to the environment. Self-specifying variant information and environment-specifying invariant information, according to Gibson, are complementary (1979, p. 183), with each reciprocally determining the other. Importantly, neither kind of complementary information is something that one is aware of. One is, however, aware of those things which the information types specify: the stable environmental layout (e.g., the non-changing proportions of a table’s rectangular surface as one approaches it) and the flowing perspective on the environment *from here* (e.g., the increasing amount of space that the table occupies in one’s visual field the closer one gets to it).

In the case of CMIG, one may worry that because the image generated and maintained is decoupled from the causal source of its spatial dimensions, there fails to be invariant information that acts as a complement to the proposed variant information. This worry may be avoided when taking a few facts about imagery into account. As we have seen above, when imagining a particular object, one’s saccadic eye patterns are similar to those which would occur if one were actually perceiving that kind of object (Thomas 1999; Fourtassi et al. 2018). This saccadic re-enacting of the object spatial relations brings with it a multiplex of visually registered stimuli. I would like to suggest that this stimuli, variant information, acts as afferent feedback for helping to guide the visual system in ways that efferently maintain (or systematically alter) the

imager's *perspective* on that which was sampled.¹⁵ In the case of CMIG, variant information does complement invariant information, but the latter need not be something that is currently coupled to for the former to continue to guide saccadic re-enactment. Any non-homogenous stimuli will serve the purpose of providing variant information (i.e., feedback) as long as it is in the particular range that does not degrade imagery. What qualifies such stimuli as variant information is *how it is used to maintain one's perspective on invariant structure in the environment*.

For example, in the disk game, by coupling to invariant spatial structure of a particular disk, one simultaneously couples to the variant, perspectival, information that accompanies it. After the disk is sampled, one's visual system decouples from it and yet remains coupled to the changing light stimulus encountered. The visual system, by using differences in light stimulus as feedback to continue the perspective-driven saccadic pattern, is able to maintain a phenomenal perspective on the spatial dimensions of the disk. This is not to say that one cannot alter one's perspective, something that is clearly the case in mental rotation tasks. The point is that extracting variant information from the environment is a process of using stimuli as feedback to drive saccades (and the visual system generally) in order to re-enact a perspective on invariant structure that is absent. This re-enactment process is not just driven by inner dynamics but guided by variant information in the environment.¹⁶

Visual variant information is generated by the interaction of saccadic patterns and differential light stimulus. Changes in impinging stimulus act as feedback—a kind of 'scaffold'—to control (confirm and correct) saccadic patterns.¹⁷ It is here that stimulus-sensitivity comes into play. Some stimuli which the visual system attempts to couple to in re-enacting perspectival structure can be of magnitudes too intense (or not intense enough) to act as feedback and thus corrupt the on-going saccadic pattern. Whilst the perceiver is simultaneously moving her sensory surfaces through the environment, high luminance stimulus that is encountered marks out the locations which saccades should avoid. Adaptive saccadic adjustments are made whilst the overall saccadic pattern is sustained.

Although variant and invariant information are complementary, it is possible for a cognizer to remain coupled to the former whilst decoupling from the latter. In the case of CMIG, one may actively couple to stimuli in the environment which allow for imagery maintenance *from here*, whilst decoupling from the invariant information that is causally dependent upon the particular object sampled. To maintain an image in CMIG is to (at least in part) adjust to the stimuli encountered in ways that allow for a lasting perspective on that which was sampled. Coupling to variant information in

¹⁵ One might object here that it is possible for one to generate a visual image in complete darkness where no such variant information is available and one's eyes nonetheless saccade in ways that reflect the spatial relations of the imagined object that was previously perceived. Assuming that one is in a condition of complete darkness—something that is rarely the case in non-experimental environments—in order for this objection to hold it must be shown that both saccades and image maintenance in the two conditions of complete darkness and in typical viewing conditions are identical. This is an open empirical question. More importantly, in the case of CMIG, typical perceptual lighting conditions are necessary.

¹⁶ I will assume that other features of imagery like colour piggyback on re-enacting perspective. (See Bompas and O'Regan 2005).

¹⁷ The generation of imagery without sampling of perceived dimensions may similarly involve the engendering of saccadic eye patterns as a result prior perceptual system biasing. These patterns are sustained (i.e., confirmed and corrected) by the generation of and coupling to variant information.

CMIG involves the tacit guiding of the perceptual system so as remain receptive to the range of stimuli that do not result in the loss of one's perspective. It is telling that the phenomenology of mental imagery—with respect to objects—is *essentially* perspectival.¹⁸ The difficulty (or more likely, impossibility) of taking more than one perspective at a time on a generated image, if this account is correct, is more to do with the fact that it is not possible to re-enact multiple perspectives simultaneously with saccades.

Although this kind of coupling does not take environment-specifying information as one of its *relata*, it is nonetheless a form of coupling to perspective-specifying, behaviour guiding variant information.¹⁹ The claim that a perceptual system may be decoupled from invariants and nonetheless operate in non-(strictly)perceptual cognition is supported by Gibson when he writes:

“...a perceptual system that has become sensitised to certain invariants (information) and can extract them from the stimulus flux can also operate without the constraint of the stimulus flux” (Gibson 1979, p. 256).

In line with Gibson's claim, the account being developed here proposes that subsequent to being sensitised to invariant structure (i.e., sampling), a perceptual system can couple to behaviour guiding variant information *sans* the invariant structure of the object sampled. Variant information is necessary for perspectival re-enactment (e.g., saccadic guidance and pupilar response) in CMIG.²⁰ It is this coupling to variant information that I shall call *variant coupling*. It may be contrasted to *invariant coupling* (i.e., coupling to invariant information).

5 Sketching a Taxonomy

Taking the proposed account of CMIG into consideration, I would now like to sketch a taxonomy of online/offline cognition which, going beyond the limited specification criteria of stimulus-present and stimulus-absent, includes weak coupling and stimulus sensitivity. The previous sections have provided reasons to think that such a taxonomy may accurately represent the underlying relations and features of CMIG. Before proceeding with this sketch, however, a few remarks are necessary. To remind the

¹⁸ I have qualified this statement to apply to objects given the possibility of generating an image that consists of nothing other than uniform colour or (i.e., a *ganzfeld*). Although there is no perspective involved in such a *ganzfeld*, generated images of objects seem to be perspectival through and through.

¹⁹ This kind of divergence of coupling to variant information from coupling to invariant information may be exemplified in the following example: whilst looking out the window of a high-speed train and staring steadily at a fixed area in the foreground, one is strongly decoupled from much of the invariant information specifying individual objects outside due to their size (say, a piece of fruit hanging from a tree) and the speed travelled. Despite this, one can couple to the present variant information specifying one's perspective; a perspective *from here* that phenomenologically encompasses one's moving at a certain velocity in relation to a static *out there*.

²⁰ Another way of interpreting Gibson here is to view imagery (or memory) as emerging from the re-enactment of the kinds of activities which were carried out when a perceiver was actually coupled to invariant information in the absence of such information (Kiverstein and Rietveld 2018). The account developed in this paper goes further than this analysis, linking re-enactment to feedback provided by coupling with variant information.

reader, the continuum thesis, beginning from the evolutionary assumption that imagery capacities developed from simpler perceptual capacities, states that perception and imagery share some of the same underlying processing. Is there a reason to believe that variant coupling names a kind of processing that is a shared feature of perception and the imagery involved in CMIG? And if so, is variant coupling a more general feature of both online and offline cognitive phenomena?

In response to the first question, the analysis of CMIG provided above suggests that endorsing the ecological claim that both variant and invariant information pick-up is involved in perception gives reason to think that weak coupling is involved in both

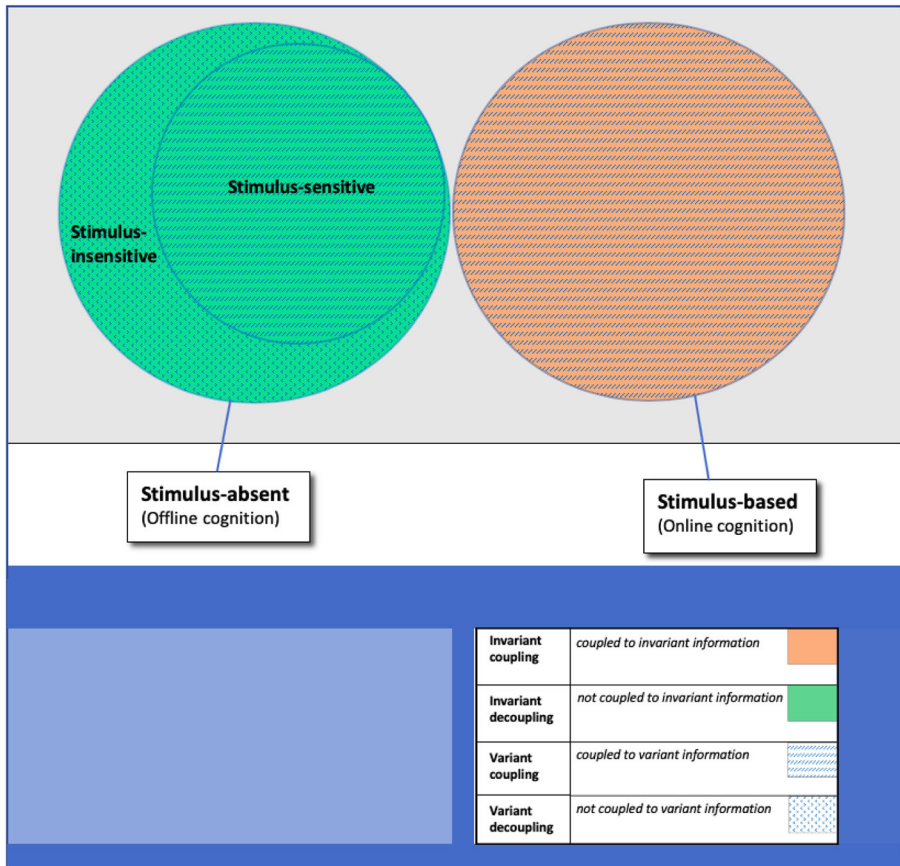


Fig. 2 Coupling in CMIG: assuming that there are offline cognitive phenomena that are stimulus-insensitive (i.e., imagery, the generation and maintenance of which is immune to environmental stimulus encountered and thus do not elicit adaptive control responses), stimulus-sensitive and stimulus-insensitive are both ways of being stimulus-absent (offline) cognition. A stimulus-sensitive process is distinguished from stimulus-insensitive process in that it is both decoupled from invariant information but coupled to variant information, whereas a stimulus-insensitive process is decoupled from both invariant information and decoupled from variant information. Stimulus-based perception is coupled to invariant information, however, like stimulus-sensitive imagery, involves also variant coupling. CMIG is a stimulus-absent process of the stimulus-sensitive type (i.e., invariant decoupling and variant coupling) which occurs *concurrently* with stimulus-based perception. This allows for variant information that is generated in stimulus-based perception to act as a feedback in stimulus absent imagery maintenance and hence to be coupled to inner-dynamics

Table 1 Implication Relations: stimulus-based implies being both variant and invariant coupling. Stimulus-sensitive implies variant coupling and invariant decoupling, and granting the possibility of stimulus-insensitive phenomena, stimulus-insensitivity implies being variant decoupling and invariant decoupling

	Variant coupling	Variant decoupling
Invariant coupling	Stimulus-based	N/A
Invariant decoupling	Stimulus-sensitive	Stimulus-insensitive

Stimulus absent =

Since stimulus-based has no determinate, there is nothing that is implied by being both decoupled to variant information and coupled to invariant information. This suggests that being coupled to invariant information is dependent upon variant coupling

perception and CMIG. In response to the latter, more general question, the claim that weak coupling is involved in all offline phenomena may certainly be challenged. The onus, however, falls on one who denies that there are stimulus-insensitive processes to demonstrate—pace our intuitions—that all offline cognitive phenomena are coupled to variant information. Any argument in favour or against stimulus-insensitive phenomena would take us too far afield, given the aim and limited scope of this paper. However, in assuming that *some* cognitive phenomena are stimulus-insensitive the following sketch of an offline/online taxonomy that does justice to CMIG begins to take shape. (See Fig. 2. and Table 1.).

In CMIG, mental images are decoupled from invariant information yet remain coupled to variant information in the environment. In contrast, online perception involves both invariant and variant coupling. Importantly, being stimulus-sensitive (or possibly stimulus-insensitive) is a way of being stimulus-absent. In other words, stimulus-sensitivity is a determinate of determinable stimulus-absent. On the other hand, there is no further determinate of stimulus-based cognition.

6 Conclusion

Both the proposed action-based account of CMIG and the taxonomical sketch on offer here may only do the work that I suggest that they may in fact do if the one is willing to concede that coupling is not restricted to the strong variety involving Gibsonian invariant information. However, the fact that variant information has been previously given a valid place within ecological analysis should make the notion of variant coupling more attractive to any theorist who is already sympathetic to the ecological framework. If the argument that I have presented above is correct, CMIG is both stimulus-absent and stimulus-sensitive: the imagery in CMIG is decoupled from the sampled invariant structure and yet remains coupled to the range of flowing variant information that enables the behavioural responses for maintenance of a perspective on that which was sampled. As such, the internal dynamics during CMIG are not detached

from the environment but subtly coupled to it as the environment provides the feedback required for a stimulus-absent perspective.

CMIG is an interesting phenomenon because it calls into question the view that all mental imagery is an offline and decoupled process. The kind of variant coupling that I have suggested underwrites CMIG is a natural extension of the ecological construct of environmental information. This deployment of the ecological framework to the analysis of CMIG has been shown to yield a novel taxonomy of online/offline cognition which goes beyond stimulus-presence and stimulus-absence. In virtue of the place this taxonomy allows stimulus sensitivity, it is able to do justice to CMIG's status as an offline and coupled phenomenon. Despite any hard to shake reluctance that one may have in accepting the accuracy of this taxonomic sketch, it may be nonetheless reasonably concluded from the case study of CMIG provided that any online/offline distinction which fails to allow for stimulus sensitive cognitive processes is inadequate. CMIG demonstrates that the online/offline bifurcation fails to be as clear-cut as it seems.

What ramifications might this analysis of CMIG and variant coupling have for the general project of articulating an ecologically-informed enactivism. Assuming that some accounts of ecological enactivism attempt to account for all basic cognition with the notion of inner and environmental dynamical coupling (Kiverstein and Rietveld 2018; Rietveld and Kiverstein 2014; Bruineberg et al. 2016/2017), one reason that the ecologically-informed enactivist might be moved to adopt the notion of variant coupling is the following thought: even in those cases in which cognizers fail to be coupled to invariant information in their environment, they may be thought to be continuously—to varying degrees—coupled to variant information in it. For the purposes of sketching an online/offline taxonomy that can account for CMIG, it was assumed that there are stimulus-insensitive phenomena. Going against this assumption, it may be asked *whether there is ever a situation in which a cognizer is completely isolated from impinging environmental stimuli?*

Answering this question negatively might suggest that all mental imagery could be stimulus-sensitive to varying degrees. This would allow the 'eco-enactivist' to work backwards from those putative cases of stimulus-absent cognition to considerations about the causal effect of variant environmental information upon various types of non-comparative mental imagery generation or offline cognition more generally. Whether or not all mental imagery is in fact stimulus-sensitive is an empirical question which conceptual analysis alone cannot decide. Even in the case that there are forms of offline cognition which are stimulus-insensitive, the concepts of stimulus-sensitivity may nonetheless provide a means for more carefully distinguishing those phenomena that are coupled in those distinct ways elaborated in this paper from those that fail to involve coupling *tout court*.

Acknowledgements I would like to thank Andy Clark, Anil Gomes, Julian Kiverstein, Becky Millar, Giovanni Pezzulo, Nina Poth and Dave Ward, in addition to two anonymous reviewers, for the many invaluable discussions and helpful comments that made this paper in its current form possible.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and

indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ashby, W.R. 1952. *Design for a brain*. 1st ed. London: Chapman & Hall.
- Ballard, D., M. Hayhoe, P. Pook, and R. Rao. 1997. Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences* 20: 723–742.
- Bickerton, D. 1996. *Language and human behaviour*. London: UCL Press.
- Binda, P., M. Pereverzeva, and S.O. Murray. 2013. Pupil constrictions to photographs of the sun. *Journal of Vision* 13 (6): 1–10. <http://www.journalofvision.org/content/13/6/8>. <https://doi.org/10.1167/13.6.8>.
- Bisiach, E., and C. Luzzatti. 1978. Unilateral neglect of representational space. *Cortex* 14: 129–133.
- Bompas, A., and J. O'Regan. 2005. More evidence for sensorimotor adaptation in color perception. *Journal of Vision* 6: 145–153.
- Bruineberg, J., J. Kiverstein, and E. Reitveld. 2016. The anticipating brain is not a scientist: The free-energy principle from an ecological-enactive perspective. *Synthese*. <https://doi.org/10.1007/s11229-016-1239-1>.
- Bruineberg, J., T. Chemero, and E. Reitveld. 2017. General ecological information supports engagement with affordances for 'higher' cognition. *Synthese*. <https://doi.org/10.1007/s11229-018-1716-9>.
- Buswell, G.T. 1935. *How people look at pictures: A study of the psychology of perception in art*. Chicago: University of Chicago Press.
- Chemero, T. 2009. *Radical embodied cognitive science*. Cambridge, MA: MIT Press.
- Cisek, P., and J. Kalaska. 2010. Neural mechanisms for interacting with a world full of action choices. *The Annual Review of Neuroscience* 33: 269–298.
- Clark, A. 1997. *Being there: Putting brain, body, and world together again*. Cambridge: MIT Press.
- Clark, A. 2016. *Surfing uncertainty: prediction, action and the embodied mind*. Oxford: University of Oxford Press.
- Degenaar, J., and E. Myin. 2014. Representation-hunger reconsidered. *Synthese*. <https://doi.org/10.1007/s11229-014-0484-4>.
- Farah, M. 2000. *The cognitive neuroscience of vision*. Wiley-Blackwell.
- Foglia, L., and F. Grush. 2011. The limitations of a purely enactive (non-representational) account of imager. *Journal of Consciousness Studies* 18 (5–6): 35–43.
- Fourtassi, M., A. Hajjioui, C. Urquizar, Y. Rossetti, G. Rode, and L. Pisella. 2018. Iterative fragmentation of cognitive maps in visual imagery tasks. *Annals of Physical and Rehabilitation Medicine*. <https://doi.org/10.1016/j.rehab.2014.03.1583>.
- Gerrans, P. 2007. Mechanisms of madness: Evolutionary psychiatry with evolutionary psychology. *Biology and Philosophy* 22: 35–56.
- Gibson, J.J. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Jost, J. 2015. Sensorimotor contingencies and the dynamical creation of structural relations underlying percepts. In *The pragmatic turn: Toward action-oriented views in cognitive science*, ed. Andreas K. Engel, Karl J. Friston, and Danica Kragic, 121–138. Cambridge, London: MIT Press.
- Keogh, R., and J. Pearson. 2011. Mental imagery and visual working memory. *PLoS One* 6 (12): e29221. <https://doi.org/10.1371/journal.pone.0029221>.
- Keogh, R., Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Vision*, 14(12):7, 1–13. doi: <https://doi.org/10.1167/14.12.7>.
- Kilpeläinen, M., and J. Theeuwes. 2016. Efficient avoidance of the penalty zone in human eye movements. *PLoS One* 11 (12). <https://doi.org/10.1371/journal.pone.0167956>.
- Kiverstein, J., and E. Reitveld. 2018. Reconceiving representation-hungry cognition: an ecological-enactive proposal. *Adaptive Behavior* 26 (4): 147–163. <https://doi.org/10.1177/1059712318772778>.
- Kosslyn, S. 1978. Measuring the visual angle of the mind's eye. *Cognitive Psychology* 10: 356–389.
- Kosslyn, S. 1980. *Image and mind*. Cambridge: Harvard University Press.
- Kosslyn, S. 2005. Mental images and the brain. *Cognitive Neuropsychology* 22 (3–4): 333–347.
- Lewis, D. 1973. Causation. *Journal of Philosophy* 70: 556–567.

- Miller, D., and L. Marks. 1992. Mental imagery and sound effects in radio commercials. *Journal of Advertising* 21: 83–93.
- Nanay, B. 2017. Multimodal mental imagery. *Cortex* 105: 125–134.
- Pezzulo, G. 2017. Tracing the roots of cognition in predictive processing. In *Philosophy and predictive processing*, ed. T. Metzinger and W. Wiese, vol. 20. Frankfurt am Main: MIND Group. <https://doi.org/10.15502/9783958573215>.
- Pezzulo, G., and P. Cisek. 2016. Navigating the affordance landscape: Feedback control as a process model of behavior and cognition. *Trends in Cognitive Science* 20 (6): 414–424.
- Rietveld, E., and J. Kiverstein. 2014. A rich landscape of affordances. *Ecological Psychology* 26 (4): 325–352. <https://doi.org/10.1080/10407413.2014.958035>.
- Shepard, R.N., and J. Metzler. 1971. Mental rotation of three-dimensional objects. *Science* 171 (3972): 701–703.
- Sherwood, R., and J. Pearson. 2010. Closing the mind's eye: Incoming luminance signals disrupt visual imagery. *PLoS One* 5 (12): e15217. <https://doi.org/10.1371/journal.pone.0015217>.
- Thomas, N. 1999. Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science* 23: 207–245.
- Varela, F., E. Thompson, and E. Rosch. 1991. *The embodied mind: Cognitive science and human experience*. MIT Press.
- Warren. 2005. Direct perception. *Philosophical Topics* 33 (1): 335–361.
- Yarbus, A. 1967. *Eye movements and vision*. New York: Plenum press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.