SI: BETWEEN VISION AND ACTION

Between vision and action: introduction to the special issue

Gabriele Ferretti[1] · Silvano Zipoli Caiani[1]

The study of the mind is quite a universe. Within this universe, we find the world of perception. Those who have landed in this world know that this is a place of wonderful and breathtaking landscapes. Among the landscapes, we can find within the world of perception, there is one we have to fear most: the jungle between vision and action. Two main kinds of adventurers are brave enough to pursue the routes in this jungle: the philosopher and the neuroscientist.

This special issue is a logbook of this trip along the insidious and marvelous pathways between vision and action.

1 The map of the philosopher

In recent years, three of the most famous philosophical maps, whose aim is to provide a description of the crucial links between vision and action, have been introduced and advocated by different philosophical theories: the *Classical Computational View*, the *Embodied View*, and the *Enactive View*.

1.1 The classical computational view (CC)

According to CC, the relation between vision and action is as a one-way route, which connects the perceptual system with the motor apparatus, but not the other way around. CC maintains that vision provides the input to cognitive processing, whereas the execution of action is the related output.

The moral is that vision and action are first and foremost peripheral events, while the central cognitive process is instantiated between the visual stimulus and the motor response (Hurley 2001).

✉ Silvano Zipoli Caiani
   silvano.zipolicaiani@unifi.it

[1] University of Florence, Florence, Italy

1.2 The embodied view (EV)

According to EV, cognitive processes are deeply dependent on the morphological properties of our bodies, as well as on the way such bodies can move, so that a clear-cut distinction between central cognition on the one hand, and perception and action on the other is very hard to outline. Accordingly, the bodily representations subserving perceptual and motor processing are inextricably linked, so that we can consider vision and action as twin sisters (Clark 1998; Lakoff and Johnson 1999; Shapiro 2010).

1.3 The enactive view (EN)

EN has recently stressed the need for removing representational items from the study of the mind, providing the debate with new arguments in support of *non-representational* theories of mind and cognition (Chemero 2011; Gallagher 2017; Hutto and Myin 2017; Varela et al. 1991).

Additionally, three main forms of interlock between vision and action have been suggested:

1. *Vision for action*: vision serves action guidance by allowing us to perceive action possibilities in the environment (Gibson 1979; Jacob and Jeannerod 2003; Jeannerod 2006; Nanay 2013; Pacherie 2011; Ferretti 2016b; Ferretti and Zipoli Caiani 2018).
2. *Vision is a form of action*: movement shapes the processing of visual information, as movement and action lead the observer to appreciate structured and familiar patterns of change in the sensory stimulation with respect to the way this movement is performed i.e. '*Sensorimotor Understanding*' (O'Regan and Noë 2001; Noë 2004). This notion is famous among *action-based theories of perception* (Briscoe and Grush 2015).
3. *Action for vision*: action processing biases and directs our visual processing: visual content is (even if sometimes only partly) determined by our intention to act, which depends on our motor capacities (De Vignemont 2018: 1.3 this issue; Nanay 2018 this issue).

Before matching these ideas with the measurement of the neuroscientist, we have to examine the latter.

2 The measurements of the neuroscientist

The best framework proposed by cognitive neuroscience over the last years is the '*Two Visual Systems Model*' (TVSM). We will briefly discuss the evolution of such a model, from the perspective of classical vision science, to the recent developments of it.

According to a '*Classical Model in Vision Science*' (CVS), the most important task of vision is object recognition, which can be used for the visual guidance of action, in a process that reflects a causal chain. The cognitive system approaches the detection of action possibilities in the environment as a serial process. First comes the visual finding of the salient properties of the targets. Then, the visual brain constructs an internal representation of the environment. Thus, the cognitive system looks for the available motor possibilities therein. Finally, it follows the computational specification of an action plan, which in turn shapes the motor commands available for suitable motor interaction (Newell 1971).

Remarkably, CVS allows measurements of the space between vision and action that reflect the map introduced by CC in the previous section. Notably, CVS and CC agree that there are no shortcuts in the route linking the input to the output, moreover, they both suggest a one-way direction, making impossible any influence of the motoric processing on the '*encapsulated*' content of (early) vision (see Brogaard and Gatzia 2017; Raftopoulos 2001, 2014).

The classical view has turned out to be only an approximation of the measurement we can carry out in the jungle. Over the past decades, indeed, the availability of new anatomo-functional evidence has inspired a groundbreaking framework within the cognitive sciences of vision. Famously, following the preliminary work by Ungerleider and Mishkin's (1982), Goodale and Milner (1995/2006) proposed what is nowadays known as the '*Two Visual Systems Model*' (TVSM). This model has changed face over the years, due to the ever-growing evidence registered. We will discuss here the initial formulation of it and the most recent one: the '*Segregation Hypothesis*' and the '*Interactionist Hypothesis*'.

2.1 The '*Segregation Hypothesis*'

The initial formulation of the TVSM suggests a new picture concerning the hodology of our visual system, by recognizing the presence of (at least) two segregated visual pathways in our visual cortex (Milner and Goodale 1995/2006). From an anatomical point of view, a ventral stream, also known as the occipito-temporal network, stretches from the primary visual cortex to the infero-temporal cortex. The dorsal stream, also known as the occipito-parietal network, goes from the primary visual cortex to the posterior parietal cortex, with specific connections to the premotor areas. From a functional perspective, the ventral stream allows primarily conscious, but also unconscious, visual object recognition, which subserves perception from an *allocentric* frame of reference. Conversely, the dorsal stream allows the (uniquely) unconscious visual guidance of action and the related attribution of action properties

to the objects we perceive, which subserves perception from an *egocentric* frame of reference, within the *peripersonal* space of the observer (Andersen and Buneo 2003; Goodale and Milner 1992; Milner and Goodale 1995; Rizzolatti and Luppino 2001).

Now, the '*Segregation Hypothesis*' establishes an *in*direct link between vision and action, and, thus, between semantic and pragmatic processing, supporting the idea of the presence of a one-way route from the visual system to the motor system. This amounts to conceive the interaction between visual and action processing as occurring in the flow of an ordered, serial, hierarchical process of a causal chain.

2.2 The '*Interactionist Hypothesis*'

Despite its great fame, the '*Segregation Hypothesis*' has been recently questioned both on philosophical (Kozuch 2015; Wu 2014; Shepherd 2015; Mole 2010; Briscoe 2009; Ferretti 2016b, c; Zipoli Caiani and Ferretti 2017; Brogaard 2011a, b; Nanay 2013) and empirical grounds (Verhoef et al. 2011; Perry et al. 2014; Wokke et al. 2014; Borra et al. 2007; Van Polanen and Davare 2015; Hoshi and Tanji 2007; Cohen et al. 2009; Ferretti and Chinellato 2019).

First, such an anatomo-functional dissociation is not very deep in healthy humans. Neurophysiology of vision suggests, indeed, that there is no rigid functional separation between the visual paths at various points in the processing (Jacob and Jeannerod 2003: p. 255; Briscoe 2009; Chinellato and Del Pobil 2016; Ferretti 2016b; Borra et al. 2007; Cisek and Kalaska 2010; Hoshi and Tanji 2007; Schenk and McIntosh 2010). Moreover, a more accurate analysis of optic ataxia and visual agnosia reveals that such impairments do not reflect dissociation in vision for action and visual recognition as strong and clear-cut as originally thought (Rossetti et al. 2003, 2005; Blangero et al. 2007), for philosophical discussion see (Briscoe 2009; Briscoe and Schwenkler 2015; Clark 2007; Ferretti 2017b).

Second, even dorsal vision-for-action can be affected by illusions (Kopiske et al. 2016; McIntosh and Schenk 2009; Briscoe 2009; Ferretti 2016b: 5.2).

Third, equating visual awareness only with ventral vision might be wrong: we do not dispose of a strong definition of consciousness, and no crucial evidence suggests that dorsal vision is totally detached from conscious processing (but see Brogaard 2011a, b: 5.5). Plus, the contribution of ventral processing to dorsal computations gives sometimes awareness to action-guiding vision (Ferretti 2016b: 5.5).

Although considered by many as being the best computational model of visual perception available, the assumption of TVSM as a real description of the cognitive functioning is not obvious. In his paper "The two visual systems hypothesis and contrastive underdetermination", Thor Grünbaum (2018, this issue of Synthese) argues that the TVSM suffers from undetermination problem systematically generated by the way certain assumptions about the informational nature of cognition are translated into experimental practice. This forces us to review our trust in this paradigm, shifting our attention from the relationship between vision and action to the more general issue regarding the relationship between behavioral evidence and computational mental models.

We can now analyze, in a more specific manner, the most important aspects of the interlock between vision and action in the light of an interdisciplinary research between philosophy and neuroscience.

3 Conscious vision and action

Vision can be conscious or unconscious. Thus, we have to consider this distinction when investigating the links between vision and action.

3.1 Consciousness in vision for action

Phenomenological, everyday life visual experience suggests that we are visually conscious of the properties of the objects we use in order to shape our movements. The task of philosophy is to understand whether this really is the case. Notably, there are three philosophical theses, related to the notion of *Conscious Vision for Action*:

1. The *Assumption of Experience-Based Control* (EBC): Conscious visual experience *can be used to guide and control* action (Clark 1999, 2001, 2007, 2009; Briscoe 2009; Briscoe and Schwenkler 2015).
2. The *Assumption of Experience-Based Selection* (EBS): Conscious visual experience *can be used to select targets for* action.
3. The *Grounding Thesis* (GT): Conscious visual experience that visually guided action *completely grounds* on conscious visual experience (Campbell 2002; Briscoe 2009).

Interestingly, if we follow the '*Segregation Hypothesis*', the TVSM supports the EBS, but not the EBC. On the one hand, the relation between ventral conscious recognition for action planning and dorsal visuomotor programming is similar to the one between a tele-assisted robotic device, the dorsal stream, which executes the appropriate interaction with the targets selected by the operator behind the device, i.e. the ventral stream. This is perfectly in line with EBS.

However, if we follow the most up to date interpretation of the TVSM, the '*Interactionist Hypothesis*', things are different, as there is room to support also the EBC. As we saw, the dorsal stream participates in object recognition, as well as in the generation of visuospatial awareness of objects presented in peripersonal space, which is crucial for *action planning* Then, also the dorsal stream is (partly) involved in the processing of selection commonly attributed only to the ventral one, when talking of the EBS (Brogaard 2011a, b; Ferretti 2018).

That said, GT cannot be supported. There are aspects of action that do not require conscious supervision. However, it seems that dissociation between conscious vision and motoric performance is due to several factors, as whether the action automatic or slow, whether it is performed by a skilled agent or by a novel one and whether it is directed to objects in the peripheral rather central portion of the visual field, as

also a discussion of cases of optic ataxia and visual agnosia suggest (Briscoe and Schwenkler 2015; Briscoe 2009; Ferretti 2017b).

Finally, asking about whether we can consciously guide action amounts to ask whether we can consciously see action possibilities in the environment or what we can call 'affordances' according to James Gibson's vocabulary (1979). It has been recently suggested that there is no crucial argument showing that we not literally see affordances, either consciously, as part of our visual phenomenology or unconsciously (Ferretti 2019). We can just say that some parts of our visual system, the dorsal visual system, register, at the subpersonal level of visual processing, geometrical patterns in the environment and use this information to generate appropriate motor commands. Other parts of our visual system, the ventral visual system, register geometrical patterns in the environment and allow us to access this information in our conscious visual phenomenology of shapes.

3.2 *Vision as a form of action* and consciousness

What about the relation between *Sensorimotor Understanding* and *Visual Consciousness*? According to the *Enactive View*, perceivers appreciate structured patterns of change in the sensory stimulation with respect to movement, i.e. '*sensorimotor contingencies*' (O'Regan and Noë 2001; Noë 2004). Such a theory does not suggest, however, that visual consciousness depends on our *Vision for Action*. Rather, vision depends on the estimation of sensory effects of movement. Therefore, even following the '*Segregation Hypothesis*', such an idea would not be in trouble.

In his paper titled "Sensorimotor Expectations and the Visual Field", Dan Cavedon-Taylor (2018, this issue of Synthese) starts from the basic idea that "Sensorimotor expectations concern how visual experience covaries with bodily movement". Rather than focusing on sensorimotor expectations, Cavedon-Taylor refers to this notion to describe 'our experience of the visual field itself and, in particular, our experience of its limits; that is, our ever-present visual sense of there being more to see, beyond what's currently within the visual field'.

Similarly, in his paper titled "Visual Acquaintance, Action and The Explanatory Gap" Thomas Raleigh (2018, this issue of Synthese) focuses on actions related to perceptual processing that are at the basis of spatial perception on the one hand, and on color perception on the other. The author suggests why "we should expect the specific nature of color phenomenology to remain less readily intelligible than the specific nature of visual spatial phenomenology".

3.3 *Action in Vision* and the issue of consciousness

While scholars have investigated the extents to which vision guides action, less effort has been devoted to the understanding of whether the way we can act determines what we perceive, especially in relation to *Visual Consciousness*.

This notion has not to be confused with the notion of sensorimotor understanding: one thing is to say that the way we can move can modulate our visual perception, by

changing our sensory stimulation, another is to say that the action we want to perform modulates *what* we perceive in the environment.

In his paper titled "Perception is not all-purpose" (this issue of Synthese), Bence Nanay offers one of the first original investigations of the notion of *Action for Vision*, by suggesting that "one's perceptual attention depends counterfactually on one's intention to perform an action (everything else being equal)", as well as that "one's perceptual processing depends counterfactually on one's perceptual attention (everything else being equal)". This seems to indicate that 'Perception is not all-purpose', as it depends, *ceteris paribus*, counterfactually on our intentions to act.

Furthermore, in his paper titled "Bodily Awareness and Novel Multisensory Features", Robert Briscoe (2018, this issue of Synthese) argues that, contrary to the thesis according to which perception resolves without remainder into their different modality-specific components, a special type of multisensory integration can give rise to perceptual experiences representing spatial features of a peculiar type which are relevant in programming.

4 The space(s) of vision and action

Vision and action operate in specific represented spaces. According to the '*segregation hypothesis*', ventral visual recognition operates in an allocentric frame of reference, while the dorsal visuomotor processing operates in an absolute, egocentric frame of reference. Recently, however, it has been suggested that vision is always perspectival with respect to any part of the perceiver's body, so that also visual recognition is always egocentric. For classical claims on this point see (Evans 1989; Campbell 2002; Clark 2007; Peacocke 1992).

Another distinction is the one between *Peripersonal space* and *Extrapersonal space* and *Vista space*. The first is about the portion of region close to the body in which objects are ready to hand. *Extrapersonal space* is the space immediately beyond the *Extrapersonal* space and *Vista space* arrives beyond 30 m (Ferretti 2016a, b, c, 2017a; De Vignemont 2018).

In her paper titled "Peripersonal perception in action", Frédérique de Vignemont (2018, this issue of Synthese) tackles the problem of defining what is peculiar of visual perception of objects falling within the peripersonal space of the observer, i.e. the space immediately surrounding the body, and which is commonly described as the space in which action takes place. Thus, de Vignemont suggests that there are "sensory and motor specificities of peripersonal perception", as those related to "emergency and the necessity to always be ready for impact". This constitutes, along with other proposals in the literature, an analysis of the peculiar relation between vision and space, when it comes to the space within our reach.

Someone has equated egocentric space to action space (Nanay 2011; Ferretti 2016a, c for a review): the ability to perform egocentric localization is the ability to interact with the object. But this definition does not capture the fact that egocentric space is just about the point of view, and the point of view modulates our visual perception even in scenarios related to *vista space* in which the object is far away (e.g. a mountain) and, thus, does not fall within our peripersonal action space. Therefore,

the possibility of egocentric localization of an object does not need the ability to perceive the object as manipulable: since, as we saw, vision is always egocentric (Briscoe 2009), but not all vision pertains to peripersonal space (as in the case of vista space), then, all peripersonal space is egocentrically encoded, but not all egocentric space is peripersonal space (Nanay 2011: pp. 468–469, footnote 8; Noë 2004: Ch. 3; Ferretti 2016a, c, 2017b). However, the ability to perceive the object as manipulable necessarily depends on its encoding within the peripersonal action space (Ferretti 2016a, b, c).

5 Representations in vision and action

Postulating privileged links between vision and action poses a problem: how can visually-structured information be directly integrated with information that is motorically structured? This issue has its origins in the classic idea that the visual stimulus and the motor output are characterized by their proprietary systems, and that an integration between them requires the mediation of an independent language under which these forms of information can be translated (Fodor 1980).

After the discovery of the dual streams of vision (Sect. 2.2), many scholars agree on a distinction between the formats of the information conveyed by the ventral and the dorsal pathways: while the ventral stream processes visual information in a way that leads to perceptual categorization and, therefore, which can lead to form perceptual beliefs, structured in a propositional manner, the results of dorsal processing cannot serve for such a task (Butterfill and Sinigaglia 2014; Ferretti 2016b, c; Jacob and Jeannerod 2003; Jeannerod 2006).

Now, the co-existence of many mental representational formats poses an issue that cannot be easily ignored. This issue concerns the integration between motoric states like motor representations and propositional states like intentions, both of which are needed in order to shape rational and accurate action performance (Burnston 2017; Butterfill and Sinigaglia 2014; Ferretti and Zipoli Caiani 2018; Mylopoulos and Pacherie 2016; Shepherd 2017).

The relation between our intentions, perception and action has been addressed by two contributions in this issue. Notably, in his paper titled "Intelligent action guidance and the use of mixed representational formats", Joshua Shepherd (2018, this issue of Synthese) deals with the problem of explaining the intelligent guidance of action, arguing that such an intelligence does not merely depend on abstract and syllogistic processes, as established by classical paradigms of practical reasoning. According to Shepherd, intelligent actions are guided through the employment of a combination of representational formats, including those constrained by the feedback outputted by sensorimotor processes.

In his paper titled "Goals and Targets: A Developmental Puzzle about Sensitivity to Others' Actions", Stephen Butterfill (2018, this issue of Synthese) explains the role of non-propositional information with respect to how motor representations function in the development of infants' sensitivity to others' actions. Butterfill argues that the fundamental social ability to track the targets and goals of actions depends on a rich mix of many kinds of processes, among which those underlying

motor abilities have a critical role. He then develops a *motor theory of goal tracking* according to which, at least for a certain class of tracking competences, the sensitivity to other' agents action goals involves the role of motor processing and representations.

5.1 Semantic biases of visuomotor processing

In recent years, the assumption of a lack of ongoing interactions between the visuomotor system and other channels of information processing has come under scrutiny: anatomical and functional evidence points toward the existence of additional pathways linking the dorsal stream to the ventral stream as well as to other associative areas. Such extra-stream interactions provide a way for semantic information to deeply and continuously influence the processes underlying visuomotor representations in the system. This led to hypothesize that memory-stored information related to visual targets can influence the motor representations of the same target, with a specific interplay with the agent's intentions.

In her paper titled "Affordances, context and sociality", Anna Borghi (2018, this issue of Synthese) provides a review of evidence showing the interaction between visuomotor processing and different cognitive paths. In her paper she starts by asking how the cognitive system selects the right course of action among the many possibilities that are offered by the environment. Borghi clearly introduces experimental results from different fields of cognitive sciences, showing how the detection and selection of action opportunities is differently modulated by physical, cultural and social information.

The role of social information in visuomotor processing has been addressed by two other contributions of this special issue. In her paper titled "Artifacts and affordances", Erica Cosentino (2018, this issue of Synthese) argues that the visual perception of artifacts may elicit different types of affordances, depending on their socially established functions, but also on the individual agent's creativity. Thus, Cosentino defends the idea of a division between *standard affordances*, which concern the common function of artifacts, and *ad-hoc affordances*, which refer to how they are individually manipulated.

Furthermore, in her article titled "Implicit Biases in Visually Guided Action", Berit Brogaard (forthcoming, this issue of Synthese) focuses on the influence of implicit knowledge and assumptions like social stereotypes and prejudices on action planning and execution in *Vision for Action*. Brogaard provides arguments to show that, although the social information shaping our action intentions can be unconscious, it may have consequences that make us accountable of social responsibility.

Finally, the presence of a high order influence on vision for action poses two more issues. The first issue concerns the very nature of the influence of contextual information on visuomotor processing: is this a case of cognitive penetration of vision? According to Pylyshyn's influential formulation, a system is cognitively penetrable if the function it computes is semantically sensitive to the agent's beliefs (Pylyshyn 1999). The investigation of the cognitive penetration of vision for action has been famously pursued by Nanay (2013). In her paper titled "Are visuomotor

representations cognitively penetrable? Biasing action-guiding vision", Josefa Toribio (2018, this issue of Synthese) starts from Nanay's account and addresses this problem by arguing that even though there are many well-established examples of top–down causal influences on the motor processing of visual information, they should not be considered genuine instances of cognitive penetration on *Vision for Action*. Toribio provides a subtle philosophical analysis of how semantic information impacts on *Vision for Action*, showing why such an interaction cannot figure as a case of cognitive penetration.

So far, we've been talking about vision and action in a real environment. But what if we talk about dreamlike environments? In her paper, titled "Dreaming of a stable world: vision and action in sleep", Melanie Rosen (2018, this issue of Synthese) has argued that current theory of vision and action must be extended to include dreams. According to Rosen, dreams are hallucinatory, visual experiences of a world, which appear so realistic that during dreams we interpret internally generated visual stimuli as externally generated movements whereas other apparent movement is interpreted as internally generated.

The second issue raised by the investigation about the role of higher cognitive states in *Vision for Action* concerns the understanding of the deep computational and representational structure of cognition. Here the problem is to account for the influence of semantic parameters on *Vision for Action*: if cognition is not content-involving, how can different ways of considering the same visual target be relevant for action execution?

In his paper titled "Intensional Biases in Affordance Perception: An Explanatory Issue for Radical Enactivism" (this issue of Synthese), Silvano Zipoli Caiani addresses this problem by reporting and discussing evidence that the categorization of visual stimuli shapes the dynamical interaction between the agent and the environment. The idea is that *Vision for Action* is deeply influenced by inten*s*ional biases, which do not allow for an extensional account of motor cognition, as stated by EV and EN.

References

Andersen, R. A., & Buneo, C. A. (2003). Sensorimotor integration in posterior parietal cortex. *Advances in Neurology, 93,* 159–177.

Blangero, A., Ota, H., Delporte, L., Revol, P., Vindras, P., Rode, G., et al. (2007). Optic ataxia is not only 'optic': Impaired spatial integration of proprioceptive information. *NeuroImage*, *36*, T61–T68.

Borghi, A. M. (2018). Affordances, context and sociality. *Synthese,* special issue "Between vision and action".

Borra, E., Belmalih, A., Calzavara, R., Gerbella, M., Murata, A., Rozzi, S., et al. (2007). Cortical connections of the macaque anterior intraparietal (AIP) area. *Cerebral Cortex*, *18*(5), 1094–1111.

Briscoe, R. (2009). Egocentric spatial representation in action and perception. *Philosophy and Phenomenological Research, 79,* 423–460.

Briscoe, R. E. (2018). Bodily awareness and novel multisensory features. *Synthese,* special issue "Between vision and action".

Briscoe, R., & Grush, R. (2015). Action-based theories of perception. In *The Stanford Encylcopedia of philosophy* (pp. 1–66).

Briscoe, R., & Schwenkler, J. (2015). Conscious vision in action. *Cognitive Science, 39*(7), 1435–1467.

Brogaard, B. (2011a). Conscious vision for action versus unconscious vision for action? *Cognitive Science, 35,* 1076–1104.

Brogaard, B. (2011b). Are there unconscious perceptual processes? *Consciousness and Cognition, 20*(2), 449–463.

Brogaard, B. (forthcoming). Implicit biases in visually guided action. *Synthese,* special issue "Between vision and action".

Brogaard, B., & Gatzia, D. E. (2017). Is color experience cognitively penetrable? *Topics in Cognitive Science, 9*(1), 193–214. https://doi.org/10.1111/tops.12221.

Burnston, D. C. (2017). Interface problems in the explanation of action. *Philosophical Explorations, 20*(2), 242–258. https://doi.org/10.1080/13869795.2017.1312504.

Butterfill, S. (2018). Goals and targets: A developmental puzzle about sensitivity to others' actions. *Synthese,* special issue "Between vision and action".

Butterfill, S. A., & Sinigaglia, C. (2014). Intention and motor representation in purposive action: Intention and motor representation in purposive action. *Philosophy and Phenomenological Research, 88*(1), 119–145. https://doi.org/10.1111/j.1933-1592.2012.00604.x.

Campbell, J. (2002). *Reference and consciousness*. Oxford: Oxford University Press.

Cavedon-Taylor, D. (2018). Sensorimotor expectations and the visual field. *Synthese,* special issue "Between vision and action".

Chemero, A. (2011). *Radical embodied cognitive science*. Cambridge: MIT Press.

Chinellato, E., & del Pobil, A. P. (2016). The neuroscience of action and perception. In *Cognitive systems monographs*: *The visual neuroscience of robotic grasping* (pp. 7–38). https://doi.org/10.1007/978-3-319-20303-4_2.

Cisek, P., & Kalaska, J. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience, 33*(1), 269–298.

Clark, A. (1998). *Being there: Putting brain, body, and world together again*. Cambridge: MIT Press.

Clark, A. (1999). Visual awareness and visuomotor action. *Journal of Consciousness Studies, 6,* 1–18.

Clark, A. (2001). Visual experience and motor action: Are the bonds too tight? *The Philosophical Review, 110,* 495–519.

Clark, A. (2007). What reaching teaches: Consciousness, control and the inner zombie. *The British Journal for the Philosophy of Science, 58*(3), 563–594.

Clark, A. (2009). Perception, action, and experience: Unraveling the golden braid. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2008.10.020.

Cohen, N., Cross, E., Tunikc, E., Grafton, S., & Culham, J. (2009). Ventral and dorsal stream contributions to the online control of immediate and delayed grasping: A TMS approach. *Neuropsychologia, 47,* 1553–1562.

Cosentino, E. (2018). Artifacts and affordances. *Synthese,* special issue "Between vision and action".

de Vignemont, F. (2018). Peripersonal perception in action. *Synthese*. Special issue: *Between vision and action.*

Evans, G. (1989). *The varieties of reference*. New York: Oxford University Press.

Ferretti, G. (2016a). Pictures, action properties and motor related effects. *Synthese, Special Issue: Neuroscience and Its Philosophy, 193*(12), 3787–3817. https://doi.org/10.1007/s11229-016-1097-x.

Ferretti, G. (2016b). Through the forest of motor representations. *Consciousness and Cognition, 43,* 177–196. https://doi.org/10.1016/j.concog.2016.05.013.

Ferretti, G. (2016c). Visual feeling of presence. *Pacific Philosophical Quarterly*. https://doi.org/10.1111/papq.12170.

Ferretti, G. (2017a). Pictures, emotions, and the dorsal/ventral account of picture perception. *Review of Philosophy and Psychology, 8*(3), 595–616. https://doi.org/10.1007/s13164-017-0330-y.

Ferretti, G. (2017b). Two visual systems in Molyneux subjects. *Phenomenology and the Cognitive Sciences, 17*(4), 643–679. https://doi.org/10.1007/s11097-017-9533-z.

Ferretti, G. (2018). The neural dynamics of seeing-in. *Erkenntnis*. https://doi.org/10.1007/s10670-018-0060-2.

Ferretti, G. (2019). Visual phenomenology versus visuomotor imagery: how can we be aware of action properties? *Synthese*. https://doi.org/10.1007/s11229-019-02282-x.

Ferretti, G. (Forthcoming). Why *Trompe l'oeils* deceive our visual experience. *The Journal of Aesthetics and Art Criticism*.

Ferretti, G., & Zipoli Caiani, S. (2018). Solving the interface problem without translation: The same format thesis. *Pacific Philosophical Quarterly*. https://doi.org/10.1111/papq.12243.

Ferretti, G., & Chinellato, E. (2019). Can our robots rely on an emotionally charged vision-for-action? An embodied model for neurorobotics. In J. Vallverdú & V. Müller (Eds.), *Blended cognition, the robotic challenge. Springer series in cognitive and neural systems* (Vol. 12). Cham: Springer.

Fodor, J. (1980). *The language of thought* (1st ed.). Cambridge, MA: Harvard University Press.

Gallagher, S. (2017). *Enactivist interventions: Rethinking the mind* (1st ed.). Oxford, UK: Oxford University Press.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Psychology Press.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*(1), 20–25.

Grünbaum, T. (2018). The two visual systems hypothesis and contrastive underdetermination. *Synthese,* special issue "Between vision and action".

Hoshi, E., & Tanji, J. (2007). Distinctions between dorsal and ventral premotor areas: Anatomical connectivity and functional properties. *Current Opinions in Neurobiology, 17*(2), 234–242.

Hurley, S. (2001). Perception and Action: Alternative Views. *Synthese, 129*(1), 3–40. https://doi.org/10.1023/A:1012643006930.

Hutto, D. D., & Myin, E. (2017). *Evolving enactivism: Basic minds meet content*. Cambridge: MIT Press.

Jacob, P., & Jeannerod, M. (2003). *Ways of seeing: The scope and limits of visual cognition*. Oxford: Oxford University Press.

Jeannerod, M. (2006). *Motor cognition: What actions tell the self*. Oxford: OUP Oxford.

Kopiske, K., Bruno, N., Hesse, K., Schenk, T., & Franz, V. H. (2016). The functional subdivision of the visual brain: Is there a real illusion effect on action? A multi-lab replication study. *Cortex, 79,* 130–152.

Kozuch, B. (2015). Dislocation, not dissociation: The neuroanatomical argument against visual experience driving motor action. *Mind and Language, 30*(5), 572–602.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. London: Basic Books.

McIntosh, R. D., & Schenk, T. (2009). Two visual streams for perception and action: Current trends. *Neuropsychologia, 47*(6), 1391–1396.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Mole, C. (2010). Illusions, demonstratives, and the zombie action hypothesis. *Mind*, *118*(472), 995–1011.

Mylopoulos, M., & Pacherie, E. (2016). Intentions and motor representations: The interface challenge. *Review of Philosophy and Psychology*. https://doi.org/10.1007/s13164-016-0311-6.

Nanay, B. (2011). Perceiving pictures. *Phenomenology and the Cognitive Sciences, 10,* 461–480.

Nanay, B. (2013). *Between perception and action*. Oxford: OUP Oxford.

Nanay, B. (2018). Perception is not all-purpose. *Synthese,* special issue "Between vision and action".

Newell, A. (1971). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

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

Pacherie, E. (2011). Nonconceptual representations for action and the limits of intentional control. *Social Psychology*, *42*(1), 67–73.

Peacocke, C. (1992). *A study of concepts*. Cambridge, MA: MIT Press.

Perry, C. J., Tahiri, A., & Fallah, M. (2014). Feature integration within and across visual streams occurs at different visual processing stages. *Journal of Vision, 14*(2), 10.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *The Behavioral and Brain Sciences, 22*(3), 341–365 **(discussion 366-423)**.

Raftopoulos, A. (2001). Is perception informationally encapsulated? The issue of the theory-ladenness of perception. *Cognitive Science, 25*(3), 423–451. https://doi.org/10.1207/s15516709cog2503_4.

Raleigh, T. (2018). Visual acquaintance, action & the explanatory gap. *Synthese,* special issue "Between vision and action".

Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron, 31*(6), 889–901.

Rosen, M. (2018). Dreaming of a stable world: Vision and action in sleep. *Synthese,* special issue "Between vision and action".

Rossetti, Y., Pisella, L., & Vighetto, A. (2003). Optic ataxia revisited: Visually guided action versus immediate visuomotor control. *Experimental Brain Research*, *153*(2), 171–179. https://doi.org/10.1007/s00221-003-1590-6.

Rossetti, Y., Revol, P., McIntosh, R., Pisella, L., Rode, G., Danckert, J., et al. (2005). Visually guided reaching: Bilateral posterior parietal lesions cause a switch from fast visuomotor to slow cognitive control. *Neuropsychologia, 43*(2), 162–177.

Schenk, T., & McIntosh, R. D. (2010). Do we have independent visual streams for perception and action? *Cognitive Neuroscience, 1*(1), 52–62. https://doi.org/10.1080/17588920903388950.

Shapiro, L. (2010). *Embodied cognition*. London, NY: Routledge.

Shepherd, J. (2015). Conscious action/zombie action. *Noûs*. https://doi.org/10.1111/nous.12086.

Shepherd, J. (2017). Skilled action and the double life of intention. *Philosophy and Phenomenological Research*. https://doi.org/10.1111/phpr.12433.

Shepherd, J. (2018). Intelligent action guidance and the use of mixed representational formats. *Synthese,* special issue "Between vision and action".

Toribio, J. (2018). Are visuomotor representations cognitively penetrable? Biasing action-guiding vision. *Synthese,* special issue "Between vision and action".

van Polanen, V., & Davare, M. (2015). Interactions between dorsal and ventral streams for controlling skilled grasp. *Neuropsychologya, 79,* 186–191.

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

Verhoef, B. E., Vogels, R., & Janssen, P. (2011). Synchronization between the end stages of the dorsal and the ventral visual stream. *Journal of Neurophysiology, 105*(5), 2030–2042.

Wokke, M. E., Scholte, H. S., & Lamme, V. A. F. (2014). Opposing dorsal/ventral stream dynamics during figure-ground segregation. *Journal of Cognitive Neuroscience, 26*(2), 365–379.

Wu, W. (2014). Against division: Consciousness, information and the visual streams. *Mind and Language, 29*(4), 383–406.

Zipoli Caiani, S. (2018). Intensional biases in affordance perception: An explanatory issue for radical enactivism. *Synthese,* special issue "Between vision and action".

Zipoli Caiani, S., & Ferretti, G. (2017). Semantic and pragmatic integration in vision for action. *Consciousness and Cognition, 48,* 40–54. https://doi.org/10.1016/j.concog.2016.10.009.