

Are visuomotor representations cognitively penetrable? Biasing action-guiding vision

Josefa Toribio
ICREA-UB^{1 2}

Abstract

Is action-guiding vision cognitively penetrable? More specifically, is the visual processing that guides our goal-directed actions sensitive to semantic information from cognitive states? This paper critically examines a recent family of arguments whose aim is to challenge a widespread and influential view in philosophy and cognitive science: the view that action-guiding vision is cognitively impenetrable. I argue, in response, that while there may very well be top-down causal influences on action-guiding vision, they should not be taken to be an instance of cognitive penetration. Assuming otherwise is to assign a computational role to the influencing states that they cannot perform. Although questions about cognitive penetrability are ultimately empirical, the issues addressed in this paper are largely philosophical. The discussion here highlights an important set of considerations that help better understand the relations between cognition, vision, and action.

Key words: action-guiding vision; cognitive penetrability; visuomotor representations; determination; bias

1. Introduction

Some well known experimental results show a dissociation between the content of visual experiences and the information that guides the subject's fine-grained sensorimotor action based on those experiences. A putative example involves the Ebbinghaus (or Titchener) illusion. On one version of this illusion, two circles of the same size are perceived as being different depending on whether they are surrounded by an annulus of smaller or larger circles. The circle surrounded by smaller circles appears (incorrectly) to be bigger than the other. However, when the illusion is implemented using 3D pieces and subjects are asked to pick up one of the central circles, their grip accurately corresponds to the actual size of the centre disc, instead of being scaled to the illusory size that the disc is experienced as having (Aglioti et al. 1995).

¹ ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain.

² Universitat de Barcelona. Department of Philosophy, Montalegre 6, Barcelona 08001, Spain.

Cases like the 3D version of the Ebbinghaus illusion—but also the Ponzo, the hollow face, or the Kanizsa compression illusions—where we are deceived by our perceptual experiences while our visually-guided actions remain unaffected by the illusion, seem to show that the information that guides the subject’s visually guided motor action—what I will henceforth call ‘visuomotor representations’ (Jacob & Jeannerod 2003)—does not capture how the world appears to the subject; it does not seem to be part of the subject’s experience.³

Evidence from neurological disorders in subjects whose visually guided action has been impaired due to severe trauma or pathological conditions brought on by injury or accidental exposure to neurotoxic substances also shows this kind of dissociation between visual experience and visuomotor representations. Two deficits are salient in this respect. One is optic ataxia, a neurological disorder that occurs when the patient can recognize objects but cannot reach them under visual guidance. The other is visual agnosia, characterized by the patients’ inability to recognize objects with which they can nevertheless interact successfully.⁴

Both types of evidence are typically considered in the empirical literature as supporting Milner and Goodale’s (1995, 2006) “dual visual system hypothesis” (see also Goodale et al. 1991). According to this view, the mechanisms and coding of information in conscious visual experience are different from, and quite independent of, the mechanisms and coding of information used to guide fine-grained online visuomotor action. On the one hand, the dorsal pathway connects the primary visual cortex (V1) to the posterior parietal cortex—the brain area for motor control—and it provides information for the guidance of skilled visuomotor action. Optic ataxics, i.e. those

³ Jeannerod (2006) offers an excellent discussion of the notion of visuomotor representation, characterized as the segment of the global representation of fast and automatic actions, such as the action of grasping, which are largely dominated by their visual input (Jeannerod, 2006, p. 5). For more recent discussion of the same notion, see e.g. Butterfill and Sinigaglia (2014), Ferretti (2016), Mylopoulos and Pacherie (2017), Shepherd (2017) and Zipoli Caiani and Ferretti (forthcoming). This last paper carefully examines the notion of visuomotor representation and its relation to higher cognitive processing.

⁴ Sometimes visual agnosics can actually recognise objects or even describe them through other senses, like sound or touch. Colours associated with certain objects can also help identification.

patients who can recognize objects without any difficulty but are unable to act upon them—e.g. grasp the objects or orient themselves towards them—in any appropriate way, seem to have, according to Milner and Goodale, a damaged dorsal stream. On the other hand, the ventral stream runs from V1 to the infero-temporal cortex—the conceptual and language processing area of the brain. It is considered responsible for the subject’s recognitional and classificatory abilities concerning visually guided action. The ventral stream—Milner and Goodale claim—suberves conscious perceptual judgment.⁵ Visual agnostics, those patients who fail to categorise their visual input, would thus be patients with a damaged ventral stream. These patients would be e.g. able to reach for a screwdriver with a beautifully calibrated grasp, even though they may very well pick it up at the wrong end since they cannot identify it as an object with a particular function. Sometimes, as in the case of patient DF, these subjects do not seem to have conscious visual experience of the shape and orientation of objects at all, yet they can, in forced choice conditions, engage in action-oriented tasks with objects in ways that match control groups of normal-sighted subjects.⁶ In the Ebbinghaus and other illusions, the dissociation between illusory conscious visual content and the fine-grained information that guides motor action is also usually explained by reference to this dual stream model.

The empirical literature does not, however, unequivocally support this interpretation. It has been suggested, for instance, that there may be rather more interaction between the two streams than the initial experiments seemed to suggest, and that the level of such interaction is task-dependent (see, e.g., Ellis et al. 1999). Pisella et al. (2000), Rossetti et al. (2003) and Rossetti et al. (2005) revisit the role of the ventral and the dorsal visual streams so as to question the assimilation of this anatomical distinction to the distinction between perception and action. They argue that there is indeed insufficient evidence to

⁵ I take perceptual judgments to be what in the literature is known as phenomenal (as opposed to epistemic) “seemings” (Brogaard, 2013, 2014; Lyons, 2005; Reiland, 2014, 2015; Tucker, 2010). They are mental episodes with a hybrid phenomenology: the products of interface systems between perception and cognition, sharing phenomenological aspects of both. See Toribio (2015; forthcoming) for a more detailed treatment of this notion.

⁶ Visual agnostics like DF can nevertheless perceive colours and textures. Inasmuch as colours could get associated to certain objects, it would be correct to say that DF has a certain residual visual experience of objects. But only in that indirect sense.

argue for this dissociation on the basis that optic ataxia appears to be a phenomenon that takes place only when subjects reach for an object into the periphery of their visual field or when there has to be some subtle correction of the reaching movement as a consequence of the target object having moved in some way. They also claim—against Milner and Goodale—that both the selection of action and the initial motor programming of heading may be carried out by the ventral stream. Milner and Goodale disagree about these results and provide further evidence to show that optic ataxics are indeed impaired at the level of initial motor behaviour, even though—they acknowledge—there may be two different subsystems operative within the dorsal stream (Milner & Goodale 2006, pp. 237–238). The philosophical discussion of the empirical literature regarding this issue has been considerable, and it will play an important role in the discussion below.⁷ Less important for my purposes here is the idea that visuomotor representations do not enter into the subject’s conscious visual awareness and hence are not part of the subject’s visual experience—an issue that has also been recently questioned and is a major point of debate both in vision science and the philosophy of cognitive science (see e.g. Briscoe 2008; Brogaard 2011a, 2011b; Kravitz et al. 2011). For my main concern here is whether or not visuomotor representations are, as a widespread and influential view has it, cognitively impenetrable (see e.g. Goodale 2011; Goodale & Milner 2004; Goodale & Wolf 2009; Jacob, 2005; Jacob & Jeannerod 2003; Milner 2008; Milner & Goodale 1995, 2008; Norman 2002, Raftopoulos 2001, 2005; Rizzolatti & Matelli 2003). This paper critically examines a recent family of arguments whose aim is to challenge this prevalent view in philosophy and cognitive science.

In response to such arguments, I suggest that, while there may very well be top-down causal influences on action-guiding vision, they do not affect the computations performed by action-guiding visual processes, in particular, dorsal visual stream computations. Such influences should thus not be taken to be an instance of cognitive penetration. Assuming otherwise is to assign a computational role to the influencing states that they cannot perform, for their content could not account for the specificity and diversity of their

⁷ See, e.g., Clark (2001; 2009), Dijkerman et al. (2009), Ferretti (2016a; 2016b; 2017; forthcoming), Jacob and Jeannerod (2003), Kozuch (2015), Mole (2009), Wu (2014), Zipoli Caiani and Ferretti (2017).

effects on goal-directed actions. I here rely on a view about the relationships between perception and cognition recently put forward by Daniel Burnston (2016) to argue that the alleged evidence in favour of the cognitive penetration of action-guiding vision is best characterized as evidence in favour of a weaker and far less controversial thesis, i.e., the view that top-down influences just potentiate or bias a set of already existent action-guiding visual processes. Although questions about cognitive penetrability are ultimately empirical, the issues addressed in this paper are largely philosophical. The discussion here highlights an important set of considerations that help better understand the relations between cognition, vision, and action.

2. Cognitive penetration

Pylyshyn's (1984, 1999) original formulation of the cognitive penetrability of perception, while discussing the cognitive *impenetrability* of early vision, remains a reference point in the contemporary discussion of this topic: "If a system is cognitively penetrable" Pylyshyn (1999, p. 343) contends "then the function it computes is sensitive, in a semantically coherent way, to the organism's goals and beliefs, that is, it can be altered in a way that bears some logical relation to what the person knows". For cognitive penetration to occur, there should be a direct causal influence of a background state A on the processing of information that leads to penetrated state B. But not any direct causal influence would do, according to this view. Such a direct causal influence would have to be a causal influence on the processing of incoming perceptual information that is not mediated by attention or by any background cognitive (including affective) state. My belief that it is time to turn the oven off, for example, may influence what I see by causing me to move my eyes in the direction of the kitchen clock—a typical case of overt attention. Such an influence, however, is not taken to be direct in the relevant sense, for my belief determines only where I direct my attention, not what I see once I am looking there. To accept that the perceptual input that we end up processing is a result of what we (overtly) attend to would make *CP* ubiquitous and uninteresting.

A different and more complicated issue is whether other types of cognitively driven attentional effects on perception amount to cognitive penetration. The effect of so-called feature/object-based attention, i.e., the attention we pay to particular features of objects such as their colour, is one such effect. It is a case of what it is known as *endogenous covert* attention, i.e., shifts among locations, features and objects, driven in a top-down fashion, while maintaining one's gaze fixed. The issue is relevant, specifically, when considering the cognitive penetrability of early vision, but it has also consequences for the evaluation of the cognitive penetrability thesis with regard to experience. Both e.g. Pylyshyn (1999) and Raftopoulos (2009) disregard this type of attentional effects—considered to be indirect causal influences that occur post-perceptually—as a sign of the cognitive penetrability of early vision processing. The idea seems to be that, for prior cognitive or affective states to have an influence on early vision processing, we have first to recognize the object we are looking at. Yet, object recognition only takes place in late vision.

It is also central to Pylyshyn's formulation of the cognitive penetrability of perception that the causal relationship between penetrating state A and penetrated state B has to be sensitive to the content of the states involved. Fiona Macpherson's by now classic migraine example illustrates the dangers of omitting this semantic coherence requirement when discussing not just early vision, but visual experience. My belief that today is the day of an important exam causes me to experience anxiety, which gives me a bad migraine, which, in turn, makes me perceive some flashing lights (Macpherson, 2012, p. 26). Even though my belief eventually causes the light flashes, the content of the belief and the content of the light-flashing experience are not semantically (i.e., logically or rationally) related. There is nothing about the content of my belief that relates to the content of my experience, and it would be a mistake to consider this a case of cognitive penetration.

A formulation of the cognitive penetrability of perception that includes the semantic coherence requirement may be considered too demanding, and there are indeed alternative characterizations in the market that skip this constraint (see e.g. Stokes, 2013).

However, there seem to be independent reasons to applaud its inclusion. One of the crucial features of cognitive states, such as beliefs and desires, is that they are related to each other semantically through complex networks of rational inference. Representations at the cognitive level are governed by standards of rationality that differ from those governing perceptual experiences. Beliefs, for instance, cannot go against our knowledge without thus signalling irrationality, while there is no irrationality, however, in our perceiving what is known to be false. The existence/absence of semantic coherence between cognitive and perceptual states would thus be a sign of the continuity/discontinuity between perception and cognition—precisely the kind of issue that motivated Pylyshyn’s discussion in the first place. A formulation of the cognitive penetrability of perception that incorporates the semantic coherence constrain will make its philosophical consequences much more pressing, if the hypothesis is true, for perception would not depend on cognition just counterfactually, but also rationally.

Daniel Burston (2016, p. 8) has recently offered a characterization of the cognitively penetrability of perception that, like Pylyshyn’s, focuses on the penetrability of perceptual processing as opposed to the penetrability of perceptual states and also posits a highly specific and direct causal influence of cognition on perception. I find it particularly useful for my purposes here. It is what he calls ‘the internal effect view’ (IEV):

A perceptual process P is penetrated if, over a specific input, it would perform a certain computation C leading to content R1 in the absence of a cognitive state, S, but performs a different computation C2, yielding content R2, when S is present, where the causal, semantic coherence, and computation conditions are met.

In what follows, the alleged cognitive penetrability of visuomotor representations will be checked against this characterization of cognitive penetrability as IEV.

3. The cognitive impenetrability of visuomotor representations

As I said earlier, it is common ground to endorse the view that action-guiding vision is cognitively impenetrable. This is typically taken to mean that visuomotor representations are not sensitive to semantic information from cognitive states in the way suggested by IEV. It is also common ground to think of visuomotor representations as representations generated by the dorsal stream. Generally, the claim is that dorsal, not ventral, vision is cognitively impenetrable. To wit, Milner & Goodale (2008, pp. 776-777):

It is the dorsal stream's job to use the current visual information about the size, shape, disposition of the object in egocentric coordinates ... to program and control the skilled movements needed to carry out the action ... To achieve this, the dorsal stream does not use the high-level perceptual representations of the object constructed by the ventral stream, but instead relies on current bottom-up information from the retina to specify the required movement parameters such as the trajectory of the reach.

Along the same lines, Raftopoulos (2009, p. 112) claims: "Top-down semantic influences from cognitive centers, do not affect ... the on-line control of action, which ... is the function served by the dorsal system". Even bolder, Raftopoulos (2017, p. 977) contends: "There is to date no evidence to support the existence of any cognitive effects on the dorsal system when it function on line to support fast action."

The main reasons for thinking of visuomotor representations as cognitive impenetrable overlap with the reasons for positing the dual visual system in the first place: evidence from neuropathologies, like visual agnosia and visual ataxia, and evidence from our motor behaviour when confronted with 3D versions of visual illusions. There is also evidence from neuroimaging, particularly from fMRI studies showing selective activation of either the dorsal or the ventral stream depending on the type of tasks subjects engage in (see e.g. Culham et al. 2003; Grill-Spector et al. 2000). Finally, there is also evidence from response latencies, i.e., the time it takes to process information in different visual areas. Although dorsal and ventral streams have similar response latencies when compared to other areas of the brain, there are significant differences between them—

with the dorsal stream being much faster than the ventral. Although different factors contribute to this difference in response latencies, the lower response latencies of the dorsal stream are typically explained by appealing to the processing being solely based on incoming information from the preceding areas (feed-forward information). When feed-back mechanisms play a modulating effect, they do it, at least initially, in the form of lateral inhibition; any top-down flow of information comes from signals transmitted from areas within early vision, thus bearing no evidence of cognitive effects (see e.g. Tovée 1994).

The imprint of cognitive impenetrability on visuomotor representations as a trademark of the dual visual system hypothesis debunked a common assumption both in philosophy and vision science—what Clark (2001, p. 496) calls the Assumption of Experience-Based Control (EBC), according to which:

Conscious visual experience presents the world to the subject in a richly textured way, a way that ... is ... especially apt for, and typically utilized in, the control and guidance of fine-tuned, real-world activity.

The idea that real-time control of action-guiding vision does not depend on the visual representations involved in our cognitive operations but are, instead, the result of a visual mechanism that is functionally and anatomically separate from the mechanism that generates our visual perception of the world thus calls for a revision of the traditional picture. Clark's own suggestion for connecting vision-guiding action to conscious visual experience is to make the relationship indirect, mediated by the use of other cognitive resources—especially attention—which put sensory systems in contact with higher-level faculties, such as working and episodic memory. If this picture is right, then perception is really geared towards recalling and reasoning and only indirectly towards action. When we say that visual experience (illusory or not) guides our behaviour (Clark 2001, p. 516):

We must unpack that notion of guidance as the high-level, intentional selection of action types and targets—we must depict it as the capacity to consciously use the perceptual array to identify goals, plan actions, and select skilled routines ... but not to control the fine detail of those sensorimotor routines themselves.

In the Ebbinghaus illusion, that the grasping of the middle circle is adjusted to their real size, not to their perceived size, would thus imply, according to this view, that visuomotor representations do not control the fine detail of the subject's action. Instead they guide action in a rather mediated way, by selecting the target circle and by choosing a type of action (grasping) to apply to it—both these choices being mediated by the use of higher-level cognitive faculties.⁸ We need to replace, Clark contends, an experience-based control of action-guiding vision for an experience-based selection model.⁹ According to the hypothesis of Experience-Based *Selection* (EBS), as Clark (2001, p. 512) labels it, “conscious visual experience presents the world to a subject in a form appropriate for the reason-and-memory-based selection of actions.”

I return to EBS in the final Section. In the rest of the paper, I would like to discuss a recent set of arguments (see e.g. Mahon & Wu 2015; but especially Nanay 2013a and Nanay 2013b) whose main conclusion is that action-guiding vision is, despite the widespread view just examined, cognitively penetrable. If it was, we may need to reconsider whether control (EBC), and not selection (EBS), is, after all, the most relevant assumption to account for the relationship between visual experience and goal-directed intentional action. I do not think we need to return to EBC, but we need first to examine the arguments and the evidence. I do so in the next Section by focusing on Nanay's (2013a and 2013b) treatment of the topic.

⁸ This may seem to imply that visuomotor representations are cognitively penetrable—inasmuch as they are influenced by the cognitive processes that drive the selection of types of action, even if in a mediated way. Yet, such cognitive processes do not causally *determine* the content of the relevant visuomotor representations. Their influence is of a different nature. The cognitive processes that drive the selection of types of action merely *bias* some already available visuomotor representations over other ones. That's what I take Clark's Experience-Based Selection hypothesis to suggest. See Section 5 below for further discussion of this crucial distinction.

⁹ Jacob and Jeannerod (2003, p. 16) follow Clark in claiming that visual perception “is not geared towards the guidance and control of action. Rather, it is geared towards the ‘selection’ of objects that can be either goals for visually guided actions or food for thought.”

4. The cognitive penetrability of action-guiding vision.

Nanay's argument takes this general form:

- P1: Action-guiding vision is cognitively impenetrable only if action-guiding vision just is dorsal vision.
- P2: Action-guiding vision is not just dorsal vision.
- C: Action-guiding vision is cognitively penetrable.

Most of Nanay's efforts go toward showing the truth of P2, which is interesting since no one ever thought that action-guiding vision was *just* dorsal vision. Yet, it is true that, as I said in the previous Section, most of the considerations put forward when claiming that action-guiding vision is cognitively impenetrable refer to dorsal processing. Nanay offers four different sets of considerations in support of P2. First, he considers evidence from neuroscience in the form of studies that suggest both the existence of a more anatomically complex set of pathways than just the dorsal and the ventral, and a larger interaction between them. Second, he claims that were we to identify visuomotor representations with dorsal processing, we would have to take sides on a rather heated debate, namely whether dorsal vision is necessarily unconscious. I do not consider the recommendation to avoid an interesting intellectual challenge to be a good reason for endorsing P2, so I will skip this point in my comments below. Third, Nanay argues, if action-guiding visual representations were just the result of dorsal information processing, there would be no action-guiding representations in sense modalities other than vision. In general, he claims, the multimodality of perception speaks against the cognitive impenetrability of action-guiding vision. The key argument is, finally, and unsurprisingly, an argument from visual illusions. Nanay briefly reviews a study which suggests that familiarity with the size of everyday objects (two brands of matches quite popular in the UK, especially in Scotland) affects our reaching and grasping motor behaviour. This, again, is interpreted as showing the cognitive penetrability of action-guiding vision (McIntosh & Lashley 2008).

In this Section, I focus on points one, three and four before addressing more generally the issue of the cognitive penetrability of action-guiding vision in the final Section.

Anatomical and functional complexity and interaction is, of course, old news. Rizzolatti & Matelli (2003) already offered a much more detailed description of the many neural pathways connecting the frontal, temporal, and parietal cortices. They hypothesized three visual subsystems, instead of just two. The ventral stream, which maintains the functionality typically ascribed to it, the ventro-dorsal pathway, specialized in the localization of objects in egocentric space and the recognition of actions made by others, and the dorso-dorsal pathway, whose function is the control of action online. More recently, Binkofski & Buxbaum (2013) reinforce this tripartite picture by suggesting that the two dorsal pathways correspond to two distinct action-guiding visual systems, what they call the Grasp and the Use systems. The dorso-dorsal or *Grasp* system processes information of properties such as shape, size and orientation of perceived objects so as to control for online prehensile actions over them. The dorso-dorsal pathway operates, Binkofski & Buxbaum suggest, independently of long-term conceptual information. The ventro-dorsal pathway or *Use* system is dedicated, by contrast, to the processing of sensorimotor information pertaining to the long-term storage of skilled actions related to familiar objects.

Further research into people with neurological disorders has also helped to re-shape the initial view of two completely isolated pathways. DF makes errors when interacting with tools, grasping them in non-functional ways, but also displaying awkward grasps when grasping neutral blocks. Patient A.T. (the most well-known optic ataxia patient) does badly at grasping if she is confronted with featureless plastic cylinders. Yet, if she grasps familiar objects, hand aperture adjustment improves considerably. In general, even for healthy subjects, we know that when we lift an object, visual size and memory cues influence our fingertip forces. All this suggests that there must be interactions between the dorso-dorsal, ventro-dorsal, and the ventral streams, even if we currently have very little knowledge about the details of such interactions (Schenk & McIntosh 2010).¹⁰

¹⁰ For further empirical discussion of the dorsal stream bifurcation see e.g. Borghi & Riggio (2015), Chinellato & del Pobil (2016), Rizzolatti & Sinigaglia (2008) and Turella & Lingnau (2014). More philosophically oriented discussions of the same topic can be found here: Brogaard (2011b), Ferretti (2016a; 2016b; 2016c), Gallese (2007).

Nanay does not mention either of these sets of studies. I bring them up to show that it is by now a well-known fact that inter-stream interactions are complex and ubiquitous and such interaction challenges the very functional independence of the isolated processing pathways. Yet, whether or not such interactions and functional fluidity amount to cognitive penetration is a different matter. Involvement of ventral stream processing in some of the online control of movement seems to suggest that there are indeed top-down influences from anterior and posterior regions of the ventral visual pathway onto the ventro-dorsal and dorsal-dorsal ones. It makes sense, from an evolutionary point of view, that the identification of objects, their affordances and functionality, and the online guiding adjustments for grasping and using them be based on a constant and coupled feedforward and feedback flow of information between different hierarchical processing regions within the visual areas. One would expect, however, that, to properly talk about cognitive penetration, the penetrating states would have to be, on the one hand, genuinely cognitive, for it is the influence of this type of states that makes the cognitive penetrability thesis philosophically interesting. Yet, much of what is going on when talking about interaction between different visual pathways is really built into low-level, automatic perceptual and motor processes with no access to our explicit knowledge. The distinctive coding of visual features involved in action-guiding vision represents nothing of the sort referred to by the relevant standard psychological notions (concept, belief, knowledge). One would also expect, on the other hand, that the causal influence of the penetrating states would have to be distinctive enough to meet the semantic and computations conditions that IEV demands, i.e., it would have to modify the functions that visuomotor mechanisms compute in such a way so as to clearly determine the motor output. Yet, it is difficult to see how the semantics of such cognitive states, i.e., their content, could be specific enough to determine such computational functions.¹¹ I return to this point in the final Section of the paper.

¹¹ Just to be clear. The contention here is that, even if ventral representations are cognitively penetrable in the traditional sense captured by IEV, the top-down modulations from ventral to dorsal processes are of a different nature and do not amount to a case of cognitive penetrability—even indirectly. To postulate some indirect mechanism of cognitive penetration of the kind illustrated by e.g. Macpherson's (2012) mechanism involving mental imagery would not be appropriate in this context. Partially because the content of the alleged penetrable and penetrating states is of a different kind so that ventral representations cannot not alter the information processed by dorsal mechanisms—something to be argued for in the final Section. Furthermore, such an indirect mechanism is (best) designed for capturing changes in phenomenal character,

What about Nanay's use of multimodality as an argument in favour of the cognitive penetrability of action-guiding vision? He makes two points about multimodality. A general point about multisensory information processing as proof of cognitive penetration and a more specific point about the absence of action-guiding representations in modalities other than vision, if action-guiding representations were just the representations generated by the dorsal stream. Let me start with the later. The distinction between ventral and dorsal streams is basically a functional distinction, even if there are indeed distinctive anatomical regions corresponding to their different functionalities. There is thus no reason to rule out, and plenty of reasons to posit, a distinctive anatomical regions with parallel functionalities for modalities other than vision even if, as it happens, in vision, the dorsal stream plays a particular functional role vis-à-vis action-guiding behaviour. Take auditory processing, for instance. There is now well-established evidence that there are indeed two streams of information processing in the auditory system: anteroventral and posterodorsal. The anteroventral stream plays an important role in the recognition and identification of sounds, while the posterodorsal stream helps with sensorimotor integration and spatial processing of sounds (see e.g. Lima et al. 2016).

My reply to Nanay's second point about multimodality requires a little bit more elaboration, although my treatment of the topic will inevitably be brief. First, using multisensory integration as a straightforward argument for cognitive penetration (in general, but also with regard to visuomotor representations) only works if we take modularity as synonymous of encapsulation and this, in turn, as synonymous of cognitive impenetrability—Fodor's traditional view and also a widespread view in the literature. On this assumption, multisensory integration amounts to absence of modularity because we identify modules with sense modalities. Hence, crossmodal interactions break down

as in e.g. experiences involving perception of colour. Yet, the phenomenological issue plays a very little, if any, role in the discussion of visuomotor representations. I do not claim that they do not have any phenomenal character, but it is not the phenomenology of such action-guiding representations that is the central issue in the present discussion, but the details of the resulting motor actions.

modularity, which amounts to cognitive penetrability. But there are at least two reasons to resist this move. First, as Burnston and Cohen (2015) suggest, it is possible and fruitful to rethink the very idea of modularity so that it comes apart from the idea of encapsulation and cognitive impenetrability, since they are conceptually independent. A way of doing this is to construe modularity, not in terms of encapsulation, but in terms of the narrowness of the range of parameters to which a process is sensitive. The key notion is that of anisotropy—as opposed to isotropy—in the sense defined by Fodor (1983). If we follow this recommendation, “what qualifies a process as modular is its not being isotropic [but anisotropic]—viz. that the range of input parameters to which its processing is sensitive is delimitable” (Burnston & Cohen 2015, pp. 133-134). Anisotropic processes, characteristic of the processes involved in perception, stand in opposition to isotropic processes, which are best illustrated by processes of belief fixation. Unlike what happens in perception, including multimodal perception, there is no restricted set of inputs that can fix, update and revise any particular belief. On this view, which Burnston and Cohen (2015) call *the integrativist view*, crossmodal integration just *is* the way perceptual processes work, and functional boundaries typically associated with modules in the old picture are here drawn *after* revealing such crossmodal interactions. Although the integrativist view does not come without problems, it illustrates something important for my argument: the need to examine the very idea of modularity vis-à-vis sense modalities before making a simple inference from crossmodal interactions to cognitive penetrability.

The second reason to resist Nanay’s move with regard to multimodality can be formulated as follows. Even if we hold on to the traditional notion of module, it is not clear at all that crossmodal interaction amounts to the form of top-down influence that properly characterizes the notion of cognitive penetration. There are certainly alternative explanations, which maintain crossmodal talk at the perceptual level in a kind of horizontal feedforward/feedback flow of information that avoids direct top-down causal influence of cognitive states (see e.g. Deroy 2015). Again, moving from multimodality to cognitive penetrability is definitely not the straightforward step that Nanay’s argument assumes.

As I said earlier, visual illusions play a central role in Nanay's final argument for P2, i.e., the argument against the idea that action-guiding vision just is dorsal vision. The key experiment (McIntosh & Lashleya 2008) suggests that familiarity with the size of everyday objects affects our reaching and grasping motor behaviour, and this is interpreted as showing the cognitive penetrability of action-guiding vision. The familiar objects used in the experiment are two famous brands of matches in the UK: Swan Vestas and Scottish Bluebell. Boxes of Swan Vestas are 79 mm x 45 mm x 13 mm, i.e. 25% larger than the boxes of Scottish Bluebell, which are 53 mm x 36 mm x 14 mm. McIntosh and Lashleya (2008) built a 0.8-scale replica of the Swan Vestas boxes and a 1.25 replica of the Scottish Bluebell boxes so that when viewed at different distances (360 mm and 450 mm), the retinal images of the Swan Vestas 0.8-scale replica, at 360 mm, is consistent with the retinal image of the original Swan Vestas box at 450 mm and the retinal image of the Scottish Bluebell 1.25-scale replica, at 450 mm, is consistent with the original Scottish Bluebell box at 360 mm. The experiment shows that, when the subjects grasp the 1.25-scale replica of the Scottish Bluebell box, their grip size is smaller than when they grasp the standard Swan Vestas (of the same size). And when they grasp the 0.8-scale replica of the Swan Vestas box, their grip size is larger than when they grasp the standard Scottish Bluebell box (of the same size). Overall, participants over-reached for the small replica of the Swan Vestas box at the near distance, and under-reached for the large replica of the Scottish Bluebell box at the far distance. McIntosh and Lashleya (2008, p. 2441) general conclusion is that "perceptual recognition routinely influences action programming." More specifically, they contend that "unless the perceptual and memory capabilities of the dorsal stream are vastly richer than previously thought, we could infer a necessary role for the ventral stream in matching current retinal input against these stored object representations" (McIntosh & Lashleya 2008, p. 2444). Nanay takes these results to be evidence of the cognitive penetrability of action-guiding vision. But, again, is this inference warranted?

First of all, not all experimental results are consistent with the view just sketched, even if the difference is subtle. A different and more recent study run by Christensen and

collaborators (2013) replicates the same experiment, involving the same brands of matchboxes, with German participants who had never seen these objects before. The question Christensen and collaborators ask is whether simply a high degree of distinctiveness, as opposed to actual prior familiarity with the brands of match-boxes, is required to affect visuomotor representations. What they found is interesting. They learned that size and distance influence the amplitude of the reaching component significantly, as in McInstosh & Lashleya (2008) study. In contrast, however, the maximum grip aperture remained largely unaffected. From this, Christensen and collaborators conclude that visual distinctiveness is sufficient to form reliable associations in short-term learning to influence reaching but not to influence grasping. Speculatively, we could perhaps say that the ventro-dorsal stream, which processes visual information towards reaching, is indeed affected by information from the ventral pathway, but not so the dorsal-dorsal stream, which processes information toward grasping. Be this as it may, what both studies seem to show is something we already knew: that the interactions between dorsal and ventral pathways are more complex than we initially acknowledged. Yet this, as I argued above, is not enough to settle the issue of the cognitive penetrability of action-guiding vision.

There is also a less speculative explanation to account for the reliable associations that result from the visual distinctiveness of the objects involved in these studies: perceptual learning. Perceptual learning involves low-level sensory plasticity and constant adjustment and re-organization of neural connectivity leading to long-term changes of a subject perceptual system, including the visuomotor system. Although it usually requires repeated exposure to the relevant stimuli, it can also result from very few experiences. Yet, as Pylyshyn acknowledges “the tuning of basic sensory sensitivity by task-specific repetition is not the same as cognitive penetration”. Perceptual learning has recently become a frequent strategy to counter alleged evidence of cognitive penetration in discussions of the role of background knowledge in the development of recognitional abilities of certain high-level properties (see e.g. Arstila 2016; Connolly 2014). As a hypothesis, it seems particularly relevant in the present context. For the effects of familiarity or distinctiveness in the above experiments seem to be guided by top-down

(ventral to dorsal) processes, yet this recurrent processing remains localized, i.e., the resulting adjustment and re-organization remains within the visual areas themselves. In other words, even if action-guiding vision is not just dorsal vision, the issue of the cognitive penetrability of action-guiding vision is not thereby settled.

5. Determination versus Bias

So far, I have granted that there are indeed complex causal connections between the different visual pathways, including top-down effects from the ventral stream, which, as we know, subserves categorization and object recognition. I have, however, denied that acknowledging these top-down effects amounts to granting the cognitive penetrability of action-guiding vision. I have questioned that crossmodal integration amounts to cognitive penetration by appealing to a different notion of modularity and also by arguing that low-level sensory plasticity and constant adjustment and re-organization of neural connectivity across classically conceived modules is standard and fairly isolated from the cognitive processes underlying higher cognitive faculties. I have suggested, more than properly argued that, in all the cases where it appears as though visuomotor representations are penetrated, the penetrating states are either not of the right kind, or their content is not the kind of content that can properly determine the functions computed by the relevant visuomotor representations. In this final Section, I unpack in more detail this last thought.

Let's go back momentarily to the results of the experiments showing the influence of familiarity (McIntosh & Lashleya 2008) on action-guiding vision. What is the precise nature of such an influence? For cognitive penetration to occur, the presence of a cognitive state (e.g. tokening of the intention "grasp that Scottish Bluebell matchbox" or "grasp that Swan Vestas matchbox") has to causally change the computation performed by the penetrated processes so as to yield a visuomotor representation with a content that is internally *determined* by the affected computed function. This is Pylyshyn's logic and semantic constrain, and what Burnston calls the *computation condition* in his IEV

formulation of the cognitive penetrability thesis (Burnston 2016, pp. 7-8). This picture requires some form of causal *determination* that preserves the content of the relevant cognitive state through any intermediary processes so as to token the specific, i.e., content-preserving, visuomotor representation. The variables responsible for the motor output must be specific enough at the level of the visuomotor representation for the motor output to match the content of the cognitive state—to match, let’s say, the content of the intention to grasp a familiar brand of matchboxes. Yet, this condition of internal determination of the content of the visuomotor representations by cognitive states demands a highly implausible deterministic binding between the highly abstract components of the allegedly penetrating cognitive state and the precise values that constitute the resulting visuomotor representations.¹²

What we find, instead, is that vision scientists, when discussing the effects of object familiarity on reach and grasp computations (or object distinctiveness on reach, but not grasp computations) formulate such an influence in much weaker terms—in terms of bias, as opposed to determination. McIntosh and Lashleya (2008, p. 2443) speak explicitly of “the biasing effect of familiar size”, reporting that they “observed that the pre-shaping of the hand was *biased* towards the familiar size of the target box” (McIntosh & Lashleya 2008, p. 2444, my emphasis). Throughout the paper, the claim seems to be that size familiarity results in a size-object association, which then *bias*es a motor response among a set of possible ones. Talk about association with distinctive size and depth displacement *bias*, as opposed to determination, is also present in Christensen et al. (2013) study.

This idea of cognition having a bias effect that increases the probability of a perceptual or motor output being selected among a set of already available ones, in an associative way, lies precisely at the heart of Burnston’s (2016) new proposal about the relationship between perception and action. Sensorimotor states consist, on this view, of the binding of represented values along both kinematic and perceptual dimensions, i.e.,

¹² Burnston (2017) calls this the “diversity/specificity problem”. The problem arises whenever we appeal to the tokening of a cognitive state, which despite having a (propositional) content that might correspond to a range of sensorimotor representations with different particular contents, is meant to explain the tokening of a specific one through a unidirectional and deterministic process of content-causation.

along both the specific movements involved in particular motor outputs and the perceptual properties of the object towards which the action is directed. Although there are certain distinctive kinematic and perceptual dimensions for each motor output, it is central to the bias view that such characteristic dimensions range over a possible set of values. When tokening of a concept occurs as part of an intention to act (e.g. tokening of SCOTTISH BLUEBELL MATCHBOX, as part of an intention to reach and grasp for a particular Scottish Bluebell matchbox), such a tokening has a biasing effect over visuomotor representations that potentiates, but does not determine, their specific values. The influence of cognitive states on sensorimotor representations thus is, on the bias view, fundamentally probabilistic as opposed to deterministic and, importantly, it is an influence that does not change the function computed by the mechanisms that generate sensorimotor representations (Burnston 2017).

Just a quick word about how this biasing view of the influence of cognitive states on sensorimotor representations compares with another, probability-based, model of perception known as the predictive coding hypothesis (see e.g. Friston 2010; Hohwy 2013). According to this hypothesis, higher-level cortical processing regions in our brain anticipate what the next perceptual input to a lower-level cortical processing region is going to be. Such predictions, also known as priors, are based on information already in place about the structure of the world and how likely it is, given such a higher-level model of the causal structure of the world, that a certain state of affairs will follow the state we are in. High-level predictions are sometimes inaccurate, i.e., the higher-level processing regions of the brain make predictive errors and have therefore to adjust so as to lessen the disagreement between the prediction and the lower-level input. In doing so, however, they encode a very detailed and large amount of information about the source of the perceptual signals that reach the lower-level cortical regions. On this view, the brain is treated as a giant Bayesian engine always trying to predict the next perceptual state based on a constant and coupled flow of information between different hierarchical processing regions and input signals. Perception, or so the slogan goes, is prediction error minimization.

There is a certain trend in recent literature to move from the idea that perception is prediction error minimization to the claim that perception is cognitive penetrable (see e.g. Hohwy 2017; Lupyan 2015). This appropriation of the predictive coding hypothesis seems to equate top-down modulation to cognitive penetrability. There are, however, good reasons to challenge this view, some of which have to do with priors being too far removed from the semantically pregnant states that the cognitive penetrability thesis demands, since they just seem to refer to the very connection strength between different computational elements at each level of information processing.¹³ The issue of whether if some version of Bayesian predictive processing is true, then perception is cognitively penetrable is a complex one that goes beyond the scope of this paper (but see e.g. Drayson forthcoming; Macpherson 2017). Yet, one thing is clear. Burnston's proposal, what he calls the External Effect View (EEV), is offered as an explanatory tool alternative to the cognitive penetrability thesis (IEV). This explanatory tool accounts perfectly for the kind of influences that we have been examining in this paper with regard to action-guiding vision in a way that allows us to acknowledge the complex relationships between different visual pathways and the influence of other sense modalities in visuomotor outcomes without having to resort to the much stronger thesis of IEV, and does so, importantly, while keeping standardly conceived cognitive states as the potential biasing representations. According to EEV (Burnston, 2016, p. 14):

Tokening of a concept as part of a cognitive state provides a bias towards any perceptual (and motor) processes associated with the concept, raising the probability that those processes will be applied to a perceptual stimulus.

The driving principle behind Burnston's EEV is the idea that cognitive and sensorimotor states have different representational structures: cognitive states are propositional while sensorimotor representations map the kinematic and perceptual dimensions involved in different action contexts (Burnston 2017). This difference prevents cognitive states from exerting a computational influence on sensorimotor representations according to the determination model sketched above. Burnston does not phrase the distinction between

¹³ The issue here is similar to the one I raised in Section 4 against Nanay's appeal to the existence of top-down influences from the ventral stream onto the ventro-dorsal and dorsal-dorsal regions in favour of the idea of cognitive penetrability of visuomotor representations.

the content of cognitive and sensorimotor states in terms of conceptual and non-conceptual content, but the fact that he speaks about a difference in *type* of content resembles the conceptual / non-conceptual framework as it is typically portrayed in the literature at the origin of the debate: non-conceptual content is not composed of concepts and is not compositional (see e.g. Heck 2000). Burnston talks indeed about cognitive representations as discrete or digital and of sensorimotor representations as analog, in Dretske's (1981) sense. This difference in types of content explains why the content of potentially penetrating cognitive states cannot be an input to the function computed by sensorimotor processes; it explains why there can be no determination of particular outputs as a function of the cognitive states' influence, and hence no cognitive penetrability. Cognitive states can and do exert a biasing influence, but on sets of already existent functional mappings. As such, this biasing influence does not change the identity of the function computed by the mechanisms responsible of relevant motor outputs (Burnston, 2016, p. 14).

There is a *prima facie* resemblance between EEV and what Fiona Macpherson (2015) calls 'cognitive penetration lite'—also presented by her as an alternative to “classic” cognitive penetration defined as in IEV. Here is how Macpherson characterizes cognitive penetration lite (Macpherson 2015, p. 346):

The idea is that there is a form of cognitive penetration which is such that, although it causes an experience with content *q* to come about and, on that occasion, the experience with content *q* probably would not have come about unless cognitive penetration had occurred, on other occasions it would be possible to have an experience that represented that *q* without cognitive penetration occurring.

Macpherson argues that, if there is something like cognitive penetration lite, then it is *compatible* with the content of the penetrated experience being non-conceptual—whether this is understood as the claim that the content of perceptual states is different in kind to the content of cognitive states, i.e., it is not composed of concepts and not compositional (the so-called 'content view), or as the claim that a subject need not possess the concepts involved in a correct characterization of the content of her perceptual experiences (the so-

called ‘state view’). I shall not address the complex issues involved in the conceptualism vs. non-conceptualism debate, nor will I discuss whether or not lite (or classic) cognitive penetration is compatible with the content of experience being non-conceptual in any of the possible readings of this latter notion. I just want to point out that, despite the initial resemblance, cognitive penetration lite and EEV are very different proposals. Cognitive penetration lite remains faithful to the idea of causal influence as determination while EEV changes the nature of such causal influence from determination to bias. Furthermore, Burnston’s motivation for introducing EEV, as an alternative to the cognitive penetrability thesis (IEV), is precisely the idea—independently defended—that cognition and perception have different kinds of content. Hence the question of the possible compatibility between EEV and non-conceptual content does not even arise. On Burnston’s picture, the dialectical move is rather the opposite: EEV appears plausible (from a philosophical point of view) because of the different nature of the content of perception (including visuomotor representations) and thought.

Whether or not EEV is the best way of accounting for the evidence examined in this paper is, of course, an empirical matter, but the conceptual resources of EEV help shape an understanding of the complex relationships between cognition of action-guiding vision that fits nicely the explanations of the observed phenomena provided by vision scientists. It is also interesting how EEV brings us back to Clark’s (2001) idea of vision-for-selection discussed in Section 3, i.e., the idea that visual experience presents the world to a subject in a form appropriate *for the selection* of action. Burnston’s External Effect View (EEV), according to which top-down processes affect action-guiding vision in the form of predictions that constrain a set of the most probable on-going motor responses based on available information about object recognition, reminds us of this idea of vision selecting among a set of already existing motor functional mappings. Abandoning the old inflated truth of vision-as-control was a step forward in the understanding of the complex relations between vision and action, especially when thinking about basic intentional actions, such as reaching and grasping. There is no need to go back.

Acknowledgements: Research for this paper was supported by the MINECO (Ministerio de Economía y Competitividad) via research grant MCINN FFI2014-51811, by the EC, Project: 675415 – DIAPHORA, H2020-MSCA-ITN-2015, and by AGAUR (Agència de Gestió d’Ajuts Universitaris i de Recerca) via research grant 2017-SGR-63.

References:

- Arstila, V. (2016). Perceptual learning explains Two candidates for cognitive penetration. *Erkenntnis* 81(6): 1151–1172.
- Aglioti, S., Goodale, M., & DeSouza, J. F. X. (1995). Size contrast illusions deceive the eye but not the hand. *Current Biology* 5: 679–685.
- Binkofski, F. & Buxbaum, L. J. (2013). Two action systems in the human brain. *Brain and Language* 127(2): 222–229.
- Borghi, A. M. & Riggio, L. (2015). Stable and variable affordances are both automatic and flexible. *Frontiers in Human Neuroscience* 9(351): 32–47. DOI: doi: 10.3389/fnhum.2015.00351.
- Briscoe, R. (2008). Another look at the two visual systems hypothesis. *Journal of Conscious Studies* 15: 35–62.
- Brogaard, B. (2011a). Are there unconscious perceptual processes? *Consciousness and Cognition* 20: 449–463.
- Brogaard, B. (2011b). Conscious vision for action versus unconscious vision for action? *Cognitive Science* 35(6): 1076–1104.
- Brogaard, B. (2013). Phenomenal seemings and sensible dogmatism. In C. Tucker (ed.), *Seemings and Justification: New Essays on Dogmatism and Phenomenal Conservatism*, pp. 270–289. Oxford: OUP.
- Brogaard, B. (2014). Seeing as a non-experiential mental state: The case from synesthesia and visual imagery. In R. Brown (ed.) *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*, pp. 377–394. Dordrecht: Springer.
- Burnston, D. and Cohen, J. (2015). Perceptual integration, modularity, and cognitive penetration. In John Zeimbekis and Athanassios Raftopoulos (eds). *The Cognitive Penetrability of Perception: New Philosophical Perspectives*. Oxford: OUP.
- Burnston, D. (2016). Cognitive penetration and the cognition-perception interface. *Synthese* DOI: 10.1007/s11229-016-1116-y
- Burnston, D. (2017). Interface problems in the explanation of action. *Philosophical*

- Explorations* 20(2): 242–258.
- Butterfill, S. A. & Sinigaglia, C. (2014). Intention and motor representation in purposive action. *Philosophy and Phenomenological Research* 88(1): 119–145.
- Chinellato, E. & del Pobil, A. (2016). *The visual neuroscience of robotic grasping. Achieving sensorimotor skills through dorsal-ventral stream integration*. Springer.
- Christensen, A., Borchers, S. & Himmelbach, M. (2013). Effects of pictorial cues on reaching depend on the distinctiveness of target objects. *PLOS One* 8(1): e54230. DOI: 10.1371/journal.pone.0054230.
- Clark, A. (2001). Visual experience and motor action: Are the bonds too tight? *Philosophical Review* 110(4): 495–519.
- Clark, A. (2009). Perception, action, and experience: Unraveling the golden braid. *Neuropsychologia* 47(6): 1460–1468.
- Connolly, K. (2014). Perceptual learning and the contents of perception. *Erkenntnis* 79(6): 1407–1418.
- Culham, J. C., Danckert, S. L., DeSouza, J. F. X., Gati, J. S., Menon, R. S & Goodale, M. A. (2003). Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Experimental Brain Research* 153: 180–189.
- Deroy, O. (2015). Multisensory perception and cognitive penetration. In John Zeimbekis and Athanassios Raftopoulos (eds). *The cognitive penetrability of perception: New philosophical perspectives*. Oxford: OUP.
- Dijkerman, H. C., McIntosh, R. D., Schindler, I., Nijboer, T. C. W., & Milner, A. D. (2009). Choosing between alternative wrist postures: Action planning needs perception. *Neuropsychologia* 47: 1476–1483.
- Drayson, Z. (forthcoming). Direct perception and the predictive mind. *Philosophical Studies*. DOI 10.1007/s11098-017-0999-x
- Dretske, A. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT Press.
- Ellis, R., Flanagan, J., & Lederman, S. (1999). The influence of visual illusions on grasp position. *Experimental Brain Research* 125: 109–114.
- Ferretti, G. (2016a). Through the forest of motor representations. *Consciousness and Cognition* 43: 177–196.
- Ferretti, G. (2016b). Visual feeling of presence. *Pacific Philosophical Quarterly* 99(S1): 112–136.
- Ferretti, G. (2016c). Pictures, action properties and motor related effects. *Synthese* 193(12): 3787–3817.
- Ferretti, G. (2017). Two visual systems in Molyneux Subjects. *Phenomenology and the Cognitive Sciences*. DOI: <https://doi.org/10.1007/s11097-017-9533-z>.
- Ferretti, G. (forthcoming). The neural dynamics of seeing-in. *Erkenntnis*.
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Friston K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience* 11 (2): 127–38.

- Gallese, V. (2007). The “conscious” dorsal stream: embodied simulation and its role in space and action conscious awareness. *Psyche* 13(1): 1–20.
- Goodale M.A. (2011). Transforming vision into action. *Vision Research* 51: 1567–1587.
- Goodale, M.A. & Milner, A.D. (2004). *Sights unseen*. Oxford: OUP.
- Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature* 349: 154–156.
- Goodale, M.A. & Wolf, M. (2009). Vision for action. In D. Dedrick & L. Trick (Eds.), *Computation, Cognition, and Pylyshyn*. Cambridge, MA: MIT Press.
- Grill-Spector, K., Kushnir, T., Hendler, T. & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience* 3: 837–843.
- Heck, R. (2000). Nonconceptual content and the “space of reasons”. *Philosophical Review* 109(4): 483–523.
- Hohwy, J. (2013). *The predictive mind*. Oxford: OUP.
- Hohwy, J. (2017). Priors in perception: top-down modulation, bayesian perceptual learning rate, and prediction error minimization. *Consciousness & Cognition* 47: 75–85.
- Jacob, P. (2005). Grasping and perceiving an object. In A. Brooks & K. Akins (Eds.), *Cognition and the Brain*. Cambridge: CUP.
- Jacob, P., & Jeannerod, M. (2003). *Ways of seeing: The scope and limits of visual cognition*. Oxford: OUP.
- Jeannerod, M. (2006). *Motor cognition. What actions tell the self*. Oxford: OUP.
- Kozuch, B. (2015). Dislocation, not dissociation: the neuroanatomical argument against visual experience driving motor action. *Mind and Language* 30(5): 572–602.
- Kravitz, D. J., Kadharbatcha S., Saleem, S., Baker, C. I. & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Neuroscience* 12: 217–230.
- Lima, C. F., Krishnan, S. & Scott, S. K. (2016). Roles of *supplementary motor areas in auditory processing and auditory image*. *Trends in Neuroscience* 39(8): 527–542.
- Lupyan, G. (2015). Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of Philosophy and Psychology* 6(4): 547–569.
- Lyons, J. (2005). Perceptual belief and nonexperiential looks. *Philosophical Perspectives* 19, pp. 237–256.
- Macpherson, F. (2012). Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism. *Philosophy and Phenomenological Research*, 84(1): 24–62.
- Macpherson, F. (2015). Cognitive penetration and nonconceptual content. In J. Zeimbekis and A. Raftopoulos (eds.) *The Cognitive Penetrability of Perception: New Philosophical Perspectives*. Oxford: OUP.

- Macpherson, F. (2017). The relationship between cognitive penetration and predictive coding. *Consciousness and Cognition* 47: 6–16.
- Mahon, Brad & Wu, Wayne (2015). Cognitive penetration of the dorsal visual stream? In John Zeimbekis and Athanassios Raftopoulos (eds). *The Cognitive Penetrability of Perception: New Philosophical Perspectives*, pp. 200–217. Oxford: OUP.
- McIntosh, R.D., & Lashleya, G. (2008). Matching boxes: Familiar size influences action programming. *Neuropsychologia* 46: 2441–2444.
- Milner, A.D. (2008). Conscious and unconscious visual processing in the human brain. In L. Weiskrantz & M. Davies (Eds.), *Frontiers of Consciousness*. Oxford: OUP.
- Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: OUP.
- Milner, A. D., & Goodale, M. A. (2006). *The visual brain in action* (2nd ed.). Oxford: OUP.
- Milner, A. D. & Goodale, M. A. (2008). Two visual systems re-visited. *Neuropsychologia* 46: 774–785.
- Mole, C. (2009). Illusions, demonstratives and the zombie action hypothesis. *Mind* 118:995–1011.
- Mylopoulos, M. & Pacherie, E. (2017). Intentions and motor representations: the interface challenge. *Review of Philosophy and Psychology* 8(2): 317–336.
- Nanay, B. (2013a). *Between Perception and Action*. Oxford: OUP.
- Nanay, B. (2013b). Is action-guiding vision cognitively impenetrable? *Proceedings of the 35th Annual Conference of the Cognitive Science Society (CogSci 2013)*. Hillsdale, NJ: Lawrence Erlbaum, pp. 1055-1060.
- Norman, J. (2002). Two Visual Systems and Two Theories of Perception. *Behavioral and Brain Sciences* 25: 73–144.
- Pisella, L., Gréa, H., Tilikete, H. C., Vighetto, A., Desmurget, M., Rode, G., Boisson, D. & Rossetti, Y. (2000). An ‘automatic pilot’ for the hand in human posterior parietal cortex: toward reinterpreting optic ataxia. *Nature Neuroscience* 3: 729–736.
- Pylyshyn, Z. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.
- Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences* 22: 341–365.
- Raftopoulos, A. (2001). Is perception informationally encapsulated? The issue of the theory-ladenness of perception. *Cognitive Science* 25: 423–451.
- Raftopoulos, A. (2005). Perceptual systems and a viable form of realism. In A. Raftopoulos (Ed.), *Cognitive Penetrability of Perception*. Hauppauge: Nova Science.
- Raftopoulos, A. (2009). *Cognition and perception: How do psychology and neural science inform philosophy?* Cambridge, MA: MIT Press.
- Raftopoulos, A. (2017). Timing time: why early vision is cognitively impenetrable. *Proceedings of the Annual Meeting of the Cognitive Science Society (COGSI, 2017)*, pp. 974–979.
- Reiland, I. (2014). On experiencing high-level properties. *American Philosophical*

- Quarterly* 51, pp. 177–187.
- Reiland, I. (2015). Experience, seemings, and evidence. *Pacific Philosophical Quarterly* 96, pp. 510–534.
- Rizzolatti, G. & Matelli, M. (2003). Two different streams form the dorsal visual system: Anatomy and functions. *Experimental Brain Research* 153: 146–157.
- Rizzolatti, G. & Sinigaglia, C. (2008). *Mirrors in the brain: how our minds share actions, emotions, and experience*. Oxford: OUP.
- Rossetti, Y., Pisella, L. & Vighetto, A. (2003). Optic ataxia revisited: visually guided action versus immediate visuomotor control. *Experimental Brain Research* 153: 171–179.
- Rossetti, Y., McIntosh, R. D., Revol, P., Pisella, L., Rode, G., Danckert, J., Tilikete, C., Dijkerman, H. C., Boisson, D., Vighetto, A., Michel, F. & Milner, A. D. (2005). Visually guided reaching: bilateral posterior parietal lesions cause a switch from fast visuomotor to slow cognitive control. *Neuropsychologia* 43: 162–177.
- Schenk, T. & McIntosh, R. D. (2010). Do we have independent visual streams for perception and action? *Cognitive Neuroscience* 1(1): 52–78.
- Shepherd, J. (2017). Skilled action and the double life of intention. *Philosophy and Phenomenological Research*. DOI: <https://doi.org/10.1111/phpr.12433>.
- Stokes, D. (2013). Cognitive penetrability of perception. *Philosophy Compass* 8(7): 646–663.
- Toribio, J. (2015). Visual experience: rich but impenetrable. *Synthese*. DOI 10.1007/s11229-015-0889-8.
- Toribio, J. (forthcoming). Visual categorization. In Brian Glenney and José Filipe Pereira da Silva (Eds.) *The Senses and the History of Philosophy*. Oxford: Routledge.
- Tovée, M. J. (1994). How fast is the speed of thought? *Neuronal Processing* 4(12): 1125–1127.
- Tucker, C. (2010). Why open-minded people should endorse dogmatism, *Philosophical Perspectives* 24, pp. 529–545.
- Turella, L. & Lingnau, A. (2014). Neural correlates of grasping. *Frontiers in Human Neuroscience* 8(686). DOI: [10.3389/fnhum.2014.00686](https://doi.org/10.3389/fnhum.2014.00686).
- Wu, W. (2014). Against division: consciousness, information and the visual systems. *Mind and Language* 29(4): 383–406.
- Zipoli Caiani, S. & Ferretti, G. (2017). Semantic and pragmatic integration in vision for action. *Consciousness and Cognition* 48: 40–54.
- Zipoli Caiani, S. & Ferretti, G. (forthcoming). Solving the interface problem without translation: the same format thesis. *Pacific Philosophical Quarterly*.