Cognitive Penetration and Informational Encapsulation: Have we been failing the module?¹

Sam Clarke
York University
spclarke@yorku.ca

Jerry Fodor deemed informational encapsulation ‘the essence’ of a system’s modularity and argued that perceptual systems are modular in precisely this respect. Nowadays, few endorse his conclusion. Standardly, this is because experimental work is seen to demonstrate the cognitive penetrability of perceptual processing, where this is widely assumed to conflict with the informational encapsulation of perceptual systems. Here, I deny the conflict. I argue that cognitive penetration need not have any straightforward bearing on (a) the informational encapsulation of each and every one of our perceptual systems, (b) the informational encapsulation of each and every one of our perceptual computations, and (c) the important consequences these results were originally expected to have for a perception-cognition border, epistemology and cognitive science. I then propose that, once these points are recognised, particularly plausible cases of cognitive penetration would actually seem to evince the informational encapsulation of perceptual systems rather than refute/problematize this.

1. Introduction

Is perception modular? Jerry Fodor thought so. In a seminal (1983) contribution he clarified the suggestion, deeming informational encapsulation ‘the essence’ of a system’s modularity (p.70). He then argued that (as a matter of empirical fact) human perceptual systems are modular in precisely this respect (p.101).

This refined earlier appeals to the functional specialisation of these systems (e.g. Marr, 1982, p.325), was expected to mark a perception-cognition border (Fodor, 1983; 1985), and was expected to have important implications for the epistemology (Fodor, 1983; 1984) and computational tractability of perception (Fodor, 1983; 2000). The thing is: nowadays, few endorse Fodor’s conclusion. Standardly, this is because experimental work is seen to somehow demonstrate the cognitive penetrability of perceptual processing. Since cognitive penetration is widely assumed to be problematic for (and perhaps incompatible with) the informational encapsulation of perceptual systems, this has prompted many to abandon the modularity of perception altogether (Block,

¹ This paper was much improved by the constructive and critical comments I received from Greyson Abid, Jake Beck, Steve Butterfill, Laurenz Casser, Martin Davies, Azener Lopez, Fiona Macpherson, Mike Martin, Ian Phillips, Jake Quilty-Dunn, Nick Shea, and audiences in Austin, Athens, Glasgow, Oxford, and Toronto. An earlier draft of this paper won the 2019 CSPE Graduate Prize for work in the philosophy of perception.
forthcoming; Churchland, 1988; Lupyan, 2015; Wu, 2013) or to propose watered-down notions of modularity that take little or no stand on informational encapsulation (Coltheart, 1999; Lyons, 2015; Ogilvie & Carruthers, 2015). For some, this amounts to a change of subject (e.g. Wilson, 2008). But even when theorists defend the encapsulation of perceptual systems, they (Fodor included) tend to do so by rejecting the suggestion that cognitive penetration actually occurs (e.g. Firestone & Scholl, 2016; Fodor, 1983; 1988; Pylyshyn, 1999; but see Quilty-Dunn, forthcoming a). Thus, while their conclusions diverge, theorists of all stripes proceed as if cognitive penetrability were problematic for, and perhaps incompatible with, the informational encapsulation of perceptual systems.

The purpose of this paper is to suggest otherwise. I will defend the claim that human perception is modular where this requires (a) the full-blown encapsulation of each and every one of our perceptual systems and (b) the full-blown encapsulation of each and every computation taking place in human perception. In spite of this, I intend to remain neutral on the state of the evidence for cognitive penetration. This is because I will be arguing that even if cognitive penetration is pervasive in the way some have claimed, it is a mistake to think this must threaten the informational encapsulation of our perceptual systems and/or the informational encapsulation of our perceptual computations. Indeed, I will be arguing that it need not even threaten the important consequences their encapsulation was expected to have for a perception-cognition border, the epistemology of perception and/or the computational tractability of perceptual processing. With these suggestions made (henceforth, my conceptual point), I will then argue for a bold claim: that, once this is recognised, plausible cases of cognitive penetration would actually seem to evince the modularity of human perception, rather than undermine or refute this (I will call this my empirical point).

To make these points, I will proceed as follows: In §2, I will introduce the claim that perception is cognitively penetrable (§2.1) and Fodor’s claim that perceptual systems are informationally encapsulated (§2.2). I will then note that the full-blown informational encapsulation of each and every one of our perceptual systems would not be problematized by cognitive penetration of (at least) one variety (§2.3). To this end, I will consider the possibility that perception is underwritten by a hierarchy of modules, wherein the outputs of lower-level systems provide inputs to perceptual modules located at higher levels of processing. On a picture of this sort, cognitive penetration could occur at the joints between modules in the hierarchy while leaving each system completely encapsulated from central cognition and the rest of the mind. Furthermore, the perceptual modifications themselves (resulting from cognitive penetration occurring at the joints between modules) would not necessarily amount to perceptual computations, the upshot being that perception (as a whole) could (in principle) remain modular, through and through, in the face of
pervasive cognitive penetration. After arguing that this would not straightforwardly undermine the interesting consequences that perceptual modularity was originally expected to have (for a perception-cognition border, computational tractability, and the epistemology of perception – §2.4), I will then turn to my empirical point: identifying a particularly plausible case of cognitive penetration (cognitive penetration mediated by mental imagery) and arguing that (if actual) this would evince a modular architecture of the aforementioned sort, rather than refute or undermine this (§3.1). To close, I will address some possible objections (§3.2) and note questions for future research (§4).

2. The Conceptual Point

Let me begin by introducing my terms.

2.1 Perception as Cognitively Penetrable

It would be relatively uncontroversial to state that perception influences cognition. For example, when I perceive a red blob in front of me, as such, this will seemingly dispose me to believe (i.e. cognise) that there is a red blob in front of me (for discussion, see Johnston, 1992, p.222; Smith, 2001, p.291; Tye, 1995, pp. 143-4). To a first approximation, the claim that ‘perception is cognitively penetrable’ amounts to the more surprising suggestion that information flows in the opposite direction: that one’s cognitive states (what one believes, desires, intends, etc.) can systematically influence one’s perception or perceptual processing.

This requires qualification. The existence of cognitive penetration is a contentious issue, and some claim its existence would call for a revolution in the way we study perception, requiring researchers to begin carefully considering subjects’ beliefs, desires and expectations when investigating their perceptual psychology (Firestone & Scholl, 2016). There are, however, many mundane ways that cognition influences perception which are unsurprising and fail to have radical consequences of this sort. Thus, in order to specify the phenomenon at issue, theorists typically hold that ‘cognitive penetration’ (as they are interested in it) would need to involve one’s cognitive states exerting a direct and semantically coherent effect on their perceptual processing (Pylyshyn, 1999). In any case, it is cognitive penetration of this sort which has typically been assumed maximally problematic for the modularity of perceptual systems.

For one’s cognitive states to exert a direct effect on one’s perceptual processing, each step in the causal process would need to occur internally to the organism and at a psychological level of description (Stokes, 2012). This rules out trivial and uninteresting cognitive influences of the following sort: Mihir intentionally moves his head, and this results in changes to the
content/character of his visual processing. Here, Mihir’s perceptual processing has been affected by one or more of his cognitive states (most obviously, an intention). In spite of this, it would be rash to assume his cognitive states have thereby ‘penetrated’ his perceptual processing in any interesting sense. This is because it remains plausible to suppose that his cognitive states have simply guided an action that has led to different objects stimulating his retinas and getting interpreted by his perceptual systems (see Firestone & Scholl, 2016, pp.3-4). As such, it is plausible to suppose that relevant steps in the causal process are occurring outside of the organism (Mihir’s perceptual processing seems to be being influenced, in part, by a change in Mihir’s sensory relation to the external world) and (not unrelatingly) it is reasonable to suppose that relevant steps in the causal process are non-psychological (Mihir’s perceptual processing is being influenced by non-psychological effects at his sensory organs). In this way, Mihir’s cognitive states seem to be influencing his perceptual processing. But, prima facie, they are only doing so indirectly.

Indirect effects of this sort are widespread, uncontroversial and carefully controlled for by perception scientists. For this reason, it is widely agreed that they are not at issue in debates over the existence of cognitive penetration. They can, however, be contrasted with conceivable effects of the following sort: Aga is looking at a coin and, by believing it to be valuable, comes to perceive the coin as bigger than she would have done otherwise (see Bruner & Goodman, 1947). Here, Aga’s belief (a cognitive state) may seem to have affected her visual processing independently of what Aga actually perceives. Consequently, it may seem unlikely that the perceptual alteration could be attributed to a simple change in sensory input (for discussion, see Pylyshyn, 1999, and commentaries). Instead, it may seem to be Aga’s belief that is influencing the way(s) in which these inputs are being interpreted by her visual systems, themselves. In this way, Aga’s cognitive states could seem to be ‘directly’ influencing her perceptual processing, in that the causal influence would be occurring internally to the organism (Aga) and at a psychological level of description. Since effects of this sort are non-obvious and would be hard to control for in psychological experiments, their existence may have important consequences for how we study and think about perception (Firestone & Scholl, 2016). For this reason, among others (see Stokes, 2013), it is direct effects of cognition on perception that researchers have been interested in when they debate the existence and implications of cognitive penetration.

The claim that cognitive penetration requires a semantically coherent effect (Pylyshyn, 1999, p.343) concerns the suggestion that the perceptual effect must also bear a non-accidental, content
respecting relation to the cognitive state that brought it about (Macpherson, 2015). This is intended to preclude cases of the following sort: Betty believes that she will fail her exam and this causes her to have a migraine where this migraine causes her visual system to represent flashing lights in the periphery of her visual field (Macpherson, 2012, p.26). Here, a belief could be said to have directly affected Betty’s perceptual processing in the above way (we might allow that the relevant stages of the causal process – however numerous – all occur within Betty’s cranium and at a psychological level of description). In spite of this, the criterion of semantic coherence would seem to prevent it from qualifying as genuine cognitive penetration. Why? Because there would not seem to be any relevant, rational, or logical relation between the content of the offending cognitive state (an impending exam) and the modified content of perception (flashing lights).

Admittedly, our assessment of these cases may require appealing to intuitions. Thus, the above point may simply amount to the claim that, intuitively, exams bear no semantic relevance, or logical connection, to flashing lights (a point some find unsatisfactory – Stokes, 2013; but see Macpherson, 2015). Regardless, the example can be contrasted with one in which the relevance seems fairly obvious. For instance, we can conceive of a situation in which Neil’s belief that ‘heart shapes tend to be coloured red’ exerts a direct influence on his visual systems’ interpretation of an orange heart shape, causing it to represent the shape as red or redder than it would have done otherwise (see Delk & Fillenbaum, 1965). Here, the logical connection between the cognitive state and the perceptual modification would seem relatively clear. Consequently, it is effects of this sort that proponents of cognitive penetration purport to identify (Lupyan, 2015; Mole, 2015; Wu, 2013) and proponents of perception’s cognitive impenetrability deny exist (Firestone & Scholl, 2016; Gross et al., 2013; Gross, 2017; Pylyshyn, 1999). So, while there is debate over the existence of cognitive penetration, it is generally agreed that cognitive penetration would occur if (and perhaps only if) an organism’s cognitive states exerted a direct and semantically coherent effect on their perceptual processing. In any case, it is cognitive penetration of this sort that has been seen to problematize the informational encapsulation of our perceptual systems. So, moving forwards, this is how I will use the term.

2 Gross (2017) goes further. He claims that a semantically coherent effect would require that the relevant cognitive state(s) stand in an evidential relation to the perceptual modification. This has been criticized for ruling out possible cases of cognitive penetration by stipulation; for instance, alleged cases of penetration by our desires (Green, forthcoming). To some extent, this dispute is interest relative: on a charitable reading, the notion of cognitive impenetrability which Gross advances could still serve to usefully mark a perception-cognition border (though perhaps not an architectural perception-cognition border) if we grant that central cognitive reasoning is, or can be, directly modulated by evidential considerations. In any case, nothing I say will depend on how this dispute is resolved; thus, I will use a weaker and more accommodating notion of semantic coherence. Gross’s more restrictive notion of semantic coherence would only strengthen the case I am making by excluding cases of alleged penetration.

3 Often ‘the cognitively penetrability of perception’ is seen to concern the cognitive penetrability of perceptual experience (see Macpherson, 2012, p.27). There, the question is whether the ‘what it’s like’ (Nagel, 1976) of perceptual experience is affected by our cognitive states in relevant ways. This is not how I am understanding matters here. Since it is unclear
2.2 Informational Encapsulation

Informational encapsulation is a different matter altogether. It pertains to a restriction on the information a system has access to in its operations. Thus, while ‘cognitive penetration’ refers to a type of freedom in the flow of information between cognition and perception (specifically, a freedom for the former to affect the latter in direct and semantically coherent ways), a system A (e.g. a visual system) is encapsulated from a system B (e.g. central cognition) insofar as there is a particular type of restriction on system A’s access to the information accessible to system B.

To appreciate the nature of this restriction, some background is helpful. When Fodor introduced the term ‘informational encapsulation’ he was interested in the fact that sensory inputs underdetermine perceptual analyses (Fodor & Pylyshyn, 1981; cf. Helmholtz, 1876). While this has not gone unchallenged (Gibson, 1979), it remains a basic tenet of mainstream perception science (see Palmer, 1999; Rescorla, 2015). And, as Fodor was keen to note, it implies that perceptual systems engage in a form of “non-demonstrative inference” (1983, p.68). Why? Because if these systems’ inputs systematically underdetermine their outputs, then the non-accidental reliability, accuracy and utility of these outputs would seem to depend on the systems making some kind of informed prediction about the causes of their inputs. In this respect, each perceptual system could be deemed a “computational mechanism that projects and confirms a certain class of hypotheses on the basis of a certain body of data” (ibid.).

This has come to be common ground between Fodor and his opponents (at least those with whom we are engaging). But note: the relevant ‘body of data’ could be more or less unbounded. In principle, it could include things known or believed by the organism and its other psychological systems (see Bruner & Postman, 1949; Halle & Stevens, 1962). Thus, it is conceivable that a visual system would interpret its inputs on the basis of (e.g.) the organism’s explicit belief that ‘light tends to come from above’ (Brown & Friston, 2012; Clark, 2015, pp.85-86), or even their beliefs about weather and astrology (Fodor, 1983, p.64).

Fodor’s proposal that perceptual systems are informationally encapsulated denies this. It concerns the suggestion that the ‘body of data’ that is accessible to a given perceptual system, and (thus) used by it when interpreting its inputs, happens to “include, in the general case, considerably less than the organism may know” (1983, p.69). It is restricted to an architecturally prescribed ‘proprietary database’ that is dedicated to that system’s operations and that system’s operations how the modularity of perception relates to perceptual experience – some modular outputs may be unconscious (Mandelbaum, 2018; Quilty-Dunn, forthcoming b), and it is possible that all are (Mandelbaum, 2019) – the important question (for our purposes) is whether an organism’s cognitive states might directly affect their perceptual processing in a semantically coherent way, whether or not this results in modifications of experience (Pylyshyn, 1999).
alone (Fodor, 1985, p.3). In this respect, a system A qualifies as ‘informationally encapsulated’ from a system B, insofar as the information it has access to is limited to that which is located inside its proprietary database, and to the extent that this proprietary database fails to include information that is accessible to system B. According to Fodor, perceptual systems are notable in that they are informationally encapsulated from the system(s) of central cognition in this respect. 4

The Muller-Lyer illusion is often used to illustrate this suggestion (see Figure 1). Here, two lines are identical in length. But this is not how they appear. Typically, if you ask someone (or, perhaps, someone who has been raised in a ‘carpentered environment’: Segall et al., 1963), they will report that the bottom line (the one with arrows pointing inwards) looks longer than the top line. While there is disagreement as to precisely why this should be (contrast: Day, 1989; Gregory, 1997; Sekuler & Erlebacher, 1971) the important point is that things continue to appear this way even when the perceiver knows that this is not so. So, even when the perceiver gets out their tape measure and confirms that the lines are identical in length (and even reflects on this fact) the illusory percept persists.

![Figure 1: The Muller-Lyer Illusion](image)

Fodor’s claim that perceptual systems are informationally encapsulated makes easy sense of this. 5 On this view, the un-interpreted light hitting the retina underdetermines the length of the lines; it is consistent with the lines having an equal length (after all, they do), but it is also compatible with the bottom line being longer (and, say, further away) than the top line (Gregory, 1997). As such, the relevant perceptual system must estimate which of these possible interpretations is most likely.

---

4 Strictly speaking, this allows for degrees of encapsulation (see Coltheart, 1999). I will keep matters simple by assuming that perceptual systems are only ‘encapsulated’ from cognition if their proprietary databases fail to include any cognitive information whatsoever. This is often assumed (e.g. Bishop, 1997; Farah, 1994; Hulme & Snowling, 1992). My point will be that even if this is granted, it does not preclude the cognitive penetrability of perception. Allowing perceptual systems some access to cognition would only strengthen my case.

5 This example is widely discussed, and Fodor’s interpretation is controversial (see McCauley and Henrich, 2006; Prinz, 2006). I am simply using the example to illustrate the Fodorian hypothesis.
But assuming that it is informationally encapsulated the system can only draw on information stored within its proprietary database to inform its estimation. And, on the conjecture that it is encapsulated from central cognition, this proprietary database fails to include the beliefs and expectations of the perceiver, including those concerning the lines’ lengths. As such, it is rendered unsurprising to find that one’s knowing the lines to be of equal length has no (apparent) bearing on the way things are perceived to be.

2.3 Must Cognitive Penetration Threaten the Encapsulation of Perceptual Systems?

Having introduced my terms, it may seem inevitable that cognitive penetration would be problematic for proponents of perceptual modularity. For them, perceptual systems are informationally encapsulated in that they interpret their inputs on the basis of prescribed proprietary databases that exclude cognitive information (e.g. the organism’s beliefs, intentions and expectations). In this respect, they are impervious to cognition. But, if cognitive penetration occurs, then cognitive states exert direct and semantically coherent effects on perceptual processing. According to many, this would prove tantamount to a refutation of the hypothesis that perceptual systems are (or could be) informationally encapsulated in the above sense (e.g. Wu, 2014). In any case, it is widely assumed that if cognitive penetration of this sort were widespread it would falsify the hypothesis that perceptual systems are informationally encapsulated modules (e.g. Block, forthcoming; Firestone & Scholl, 2016; Prinz, 2006; Pylyshyn, 1999; Robbins, 2009). At best, we would be left with “scattered islands of modularity” (Prinz, 2006, p. 22).

This is a mistake. Admittedly, there is an obvious way in which cognitive penetration might occur that would undermine (or refute) the view that perceptual systems are (completely) encapsulated from cognition: If perceptual systems were to access our cognitive states and draw on these when interpreting their inputs, then perception would seem to be systematically affected by cognition in direct and (let us suppose) semantically coherent ways. Thus construed, cognitive penetration would, indeed, call into question the perceptual modularity hypothesis, since it would amount to finding that cognitive states enter into the bodies of information (the proprietary databases) that inform our perceptual systems’ computations. But this is just one possibility and, as it so happens, there are other ways in which cognitive penetration could occur. At least one of these would be completely irrelevant to each and every perceptual system’s informational encapsulation from central cognition and the rest of the mind, no matter how pervasive it might be.

To see this, note that the encapsulation of our perceptual systems does not require that each sensory modality be composed of a single encapsulated module. It is consistent with each sensory
modality comprising of a hierarchy of encapsulated systems. This is conceivable because (as we have seen) a system’s informational encapsulation pertains to a restriction on the proprietary database that informs it in the interpretation of its inputs. So, provided that the hierarchy could be identified as perceptual (by appeal to, say, function [see Marr, 1982]) and each system in the hierarchy were to process its inputs on the basis of its own proprietary database (and this alone), the sensory modality would still be entirely comprised of entirely encapsulated systems, even if the outputs of systems at one level were taken as inputs by systems at the next level of processing.

On this view we could allow that the outputs of a module, located at one level in the hierarchy, be modified by, or supplemented by the outputs of, some cognitive process before serving as inputs to the modules located at higher levels in the hierarchy. Provided that this simply occurred at the joints between modules, and did not add to the bodies of information guiding these systems in the interpretation of their inputs, it would leave the encapsulation of each system intact (i.e. it would not bear on the claim that their inputs are processed on the basis of prescribed bodies of information – proprietary databases – that fail to include cognitive information). But since these effects would occur within a functionally defined perceptual hierarchy (and would, presumably, lead to downstream effects of a semantically coherent nature) it would be accurate to speak of perception being cognitively penetrated. Thus, on a picture of this sort, cognitive penetration would be consistent with the full-blown informational encapsulation of each and every perceptual system, no matter how pervasive it might be.

One might object that this would still imply the existence of unencapsulated computational steps in the perceptual process. This might seem inevitable because the input-output function of the modification/supplementation computation (occurring at the joints between modules) would be being mediated by access to cognitive information located outside of perception. But note: this would only be problematic for the modularist insofar as it would be apt to view these modifications/supplementations as perceptual computations. And, in at least certain cases, this would plainly be inappropriate. For example: much has been made of the apparent fact that the intentional formation of a mental image involves cognitive influences on our perceptual machinery (see §3). However, a key reason why this has intrigued philosophers and cognitive scientists is the manifest fact that intentionally forming and manipulating mental images seems not to be a perceptual process (Kosslyn et al., 2006). Rather, it appears to be a kind of imagistic thought or cognition (Gauker, 2011). So, provided that the modification or supplementation of modular outputs was like this (i.e. a non-perceptual, cognitive process), and such influences failed to impact our perceptual processes except by modifying/supplementing the representations being passed between encapsulated modules in a hierarchy of the above sort, then effects of this sort would fail to put pressure on the
suggestion that every single perceptual computation is carried out by an informationally encapsulated, modular system. In this way, perceptual processing could remain an entirely modular affair, involving nothing more than computations performed by informationally encapsulated systems, despite the cognitive penetrability of perception.

Proponents of Fodorian encapsulation have done much to obscure this point. Leading proponents of the view (Firestone & Scholl, 2016, p.3; Fodor, 1983, p.64; Pylyshyn, 2003, pp.364-6) have often spoken as if there is a single module corresponding to each Aristotelian sense (thus: a single visual module with a dedicated proprietary database, a single auditory module with a dedicated proprietary database, and so forth) plus or minus one for speech perception. This invites an equivocation between the claim that a sensory modality, like vision, is cognitively impenetrable and the claim that a given module or computation sub-serving this modality is cognitively impenetrable. But, given the above, we should recognise that strictly speaking this is a mistake: it is only the latter claim that the modularity of each perceptual system (and the idea that each perceptual computation is carried out by a modular system) implies or requires.

Fodor should have acknowledged this. While he frequently spoke as though humans possess a single visual module, a single auditory module, and so forth, his exemplar of a modular process was speech perception (1983, p.64). For him, the speech module was encapsulated from both central cognition and the sensory modules realising vision, audition, and so forth. But this did not involve it receiving its inputs from private sensory transducers. Rather, the suggestion was that it would take as inputs the outputs of prescribed sensory modules – e.g. visual and auditory modules (McGurk & McDonald, 1976) – and would process these on the basis of its own proprietary database (see Liberman et al., 1967, who Fodor cites with approval; see also Liberman & Mattingly, 1985, who cite Fodor with approval). So, while Fodor and his sympathisers have often suggested that sense modalities, like vision, are each composed of a single module, they should accept as conceivable the suggestion that perception (within a given modality) be subserved by a hierarchy of modules, allowing that a subset of these provide inputs to other modules in the hierarchy (and, ultimately, central cognition). On this view, these inputs could be supplemented by the outputs of the organism’s cognitive processing, resulting in direct and semantically coherent effects of cognition on perception, at the points where representations pass between modules in the hierarchy. But, critically, this would not compromise the suggestion that each of these systems be completely encapsulated from all others, and it would not compromise the suggestion that every perceptual computation be carried out by a modular system. Thus, to reiterate, cognitive penetration of this sort would be consistent with the claim that perceptual processing is underpinned by nothing more than fully encapsulated modules.
2.4 Does This Change the Subject?

Cognitive penetration (of the above sort) would be consistent with the full-blown encapsulation of each and every one of our perceptual systems, and each and every one of our perceptual computations, no matter how pervasive it might be. Nevertheless, one might object that, in reconciling these phenomena, I have left the modularity hypothesis without the significance it was meant to have for cognitive science and the philosophy of perception. Or worse, that I have simply changed the subject. Before moving on, I wish to address these worries. To this end, I wish to consider the three most prominent reasons why the modularity of perception was originally expected to matter when Fodor introduced his hypothesis. I will argue that the above proposal has little or no impact on these and, thus, leaves the modularity hypothesis more or less as interesting/boring as it originally was.

An initial reason why the modularity of perceptual systems was originally expected to matter was for its ability to mark a perception-cognition border, thereby vindicating a folk distinction between perception and thought (Fodor, 1983, p.70; p.101; 1985, p.3). For Fodor, this stemmed from the suggestion that central cognition is ‘isotropic’ and ‘Quinean’; isotropic in that anything one knows or believes can be brought to bear in central cognitive processing (1983, p.105) and Quinean in that central cognitive inferences are sensitive to the organism’s entire web of belief (pp.107-8). Both claims are, of course, controversial (see Lewis, 1982; Mandelbaum, 2016; Norby, 2014). But, putting them together, Fodor held that (in principle) anything one knows or believes could be recognised as relevant to their reasoning on a given matter and brought to bear as a premise therein. Since this was tantamount to saying that such reasoning does not operate on the basis of a prescribed and architecturally determined proprietary database it amounted to a rejection of central cognition’s informational encapsulation. But, since cognitive penetration of the above sort leaves the claim that perceptual systems operate on the basis of prescribed proprietary databases unscathed, it allows that perceptual systems remain informationally encapsulated in a way that central cognition (as Fodor characterised it) is not. Thus, the proposal sketched in §2.3 should do nothing to undermine the thought that informational encapsulation marks perception out from central cognition, as Fodor understood matters. In this respect, perceptual modularity would seem to retain its significance.

A second reason why the informational encapsulation of perceptual systems was originally expected to matter was for its enabling perceptual processes to remain computationally tractable (Fodor, 1983, p.128). Since encapsulated systems interpret their inputs on the basis of dedicated proprietary databases, they need only consider a constrained amount of information when
interpreting their inputs. As such, it is relatively easy to see how they might sift through all of the information available to them, or organise it, in a timely manner, so as to appreciate how it bears on the interpretation of their inputs (see Ford & Pylyshyn, 1994). But note: since this \emph{(again)} derives from these systems’ operating on the basis of constrained proprietary databases – i.e. that which my proposal leaves intact – informational encapsulation would \emph{(again)} seem to retain its significance on the above proposal. Therefore, in this respect too, perceptual modularity would seem to retain the significance it was originally expected to have.

A third reason why the modularity of perception was seen to matter was for its epistemic implications (Fodor, 1984). From my perspective, this is a more difficult case. Cognitive penetration of the above sort does raise the worry that perceptual justification might (sometimes) be circular (Siegel, 2012) and/or liable to influence the reliability of our perceptual faculties (Lyons, 2011). Nevertheless, it would be wrong to dismiss my proposal as a change of subject on these grounds.

For Fodor (1984), informational encapsulation bore epistemic significance because it enabled perceptual systems to provide a ‘theory-neutral’ foundation for empirical knowledge. This was meant to derive from the fact that informationally encapsulated perceptual systems interpret their inputs on the basis of proprietary databases, free of the organism’s beliefs and expectations. But this always allowed that the organism might wilfully affect the inputs these systems receive (after all, Fodor always acknowledged that a suitably motivated subject might freely avert their eyes or intentionally shift their attention). Therefore, while cognitive penetration (of the above sort) may introduce certain kinds of epistemic worry (see above), it is unclear that an informational encapsulated system’s operations would not remain ‘theory-neutral’ in much the way Fodor claimed. After all, it would leave each system’s proprietary database free of background beliefs and expectations. While this may fall short of enabling the foundation for empirical knowledge Fodor had intended, it would retain significant epistemic import (for instance, by constraining imagistic thought in helpful ways [cf. Kind, 2016]).

Of course, Fodor’s claims about theory neutrality were always in dispute. Churchland (1988) was quick to object that an informationally encapsulated module would still process its inputs on the basis of a theory, and one which could be just as distorting of the facts as any other; its encapsulation would simply ensure that this theorising remain immune to certain kinds of rational revision. But note: if this is correct, it simply shows that informational encapsulation never really had the epistemic implications Fodor intended. And, once again, this would be true, irrespective of the above proposal.
Given disputes of this nature, contemporary proponents of perceptual modularity tend to distance their hypothesis that perception is modular from specific claims about perceptual epistemology (e.g. Lyons, 2015; Quilty-Dunn, forthcoming). To be clear: this does not commit them to the idea that there are no epistemic implications of the hypothesis. It simply underscores the fact that any epistemic upshots are just this; they are upshots of an independent architectural proposal. Thus, we can debate the informational encapsulation of our perceptual systems and, in addition, we can debate the implications this would have for the epistemology of perception. The important point (for our purposes) is that these debates are distinct. Consequently: when we examine the implications that perceptual modularity was originally expected to have when Fodor introduced his hypothesis we find little reason for thinking that the proposal sketched in §2.3 robs this of its original significance or that it simply constitutes a change of subject. Instead, we should embrace the conclusion that the informational encapsulation of each and every perceptual system and each and every perceptual computation is consistent with at least certain forms of cognitive penetration.

3. **The Empirical Point**

The preceding discussion makes a conceptual point; that (as a matter of conceptual fact) one form of cognitive penetration would be unproblematic for the full-blown informational encapsulation of each and every perceptual system and for the claim that each and every perceptual computation is carried out by an informationally encapsulated module. But this simply charts logical space – it does not tell us anything about how human minds are actually structured. Thus, one might concede the above as a conceptual possibility but doubt that real-life cases of cognitive penetration (assuming they exist) will prove consistent with the informational encapsulation of perceptual systems, let alone the claim that every perceptual computation is carried out by an informationally encapsulated module.

Turning to my empirical point, I will now suggest otherwise. Once we recognise the above, I propose that particularly plausible cases of cognitive penetration are both consistent with the informational encapsulation and modularity of our perceptual systems and, more strongly, that they actually seem to evince this.

Of course, an enormous number of empirical findings have been seen (by some) to demonstrate perception’s cognitive penetrability. It would be beyond the scope of this paper to consider each of these in turn. Thankfully, this will not be necessary to make my point. In recent debates, mental imagery has been seen to provide a particularly plausible illustration of perception’s cognitive

---

6 Yale’s Perception and Cognition Lab lists 182 independent studies, published in 42 different journals, since 1995, purporting to do just this (see [http://perception.yale.edu/Brian/refGuides/TopDown.html](http://perception.yale.edu/Brian/refGuides/TopDown.html)).
penetrability. For instance, in an influential critique of the modularity hypothesis, Jesse Prinz (2006) claims that mental imagery is “the most obvious” example of cognitive penetration, while (in a similar vein) Ned Block (2016) has described it as “the most dramatic of the known top-down effects of cognition on perception” (p.21; see also: Block, forthcoming). This is something that both theorists then take to undermine or refute the informational encapsulation and modularity of perceptual systems (see also: Dijekstra et al, 2019; Howe & Carter, 2016; Kosslyn et al., 2006). So: while we have already touched upon the possibility that cases of this sort might prove consistent with perceptual modularity if it could be shown that these (alleged) cases of cognitive penetration are, in light of my conceptual point, actually consistent with, and even indicative of our perceptual systems’ full-blown informational encapsulation, this would be a significant boon to proponents of perceptual modularity.

3.1 Mental Imagery

Several lines of evidence have been advanced in support of the suggestion that mental imagery involves or constitutes cognitive penetration. Some appeal to neural findings suggesting that perception and cognitively driven mental images recruit overlapping brain regions (Block, forthcoming; Dijekstra et al., 2019; Howe & Carter, 2016; Kosslyn et al., 2001). Prima facie, this may lend credence to the suggestion that mental imagery can (or does) involve the cognitive hijacking of perceptual resources. But, alone, it is unconvincing. For one thing, proponents of cognitive impenetrability have long been aware of these overlapping neural regions and deemed them irrelevant on the grounds that one cannot draw straightforward conclusions about the cognitive level from findings pitched at the level of neural implementation (e.g. Pylyshyn, 2002).

Additional evidence comes from behavioural studies, however. Here, proponents of cognitive penetration find evidence that cognitively driven mental images can become integrated into genuine perceptual processing (Block, 2016; Kosslyn et al., 2006; Perky, 1910). For instance, in an experiment emphasised by Block (2016), Brockmole et al. (2002) gave subjects a “locate the missing dot task”, requiring them to identify a single missing dot from a 5x5 array. Previous work had shown that if a segment from these arrays (containing 12 dots) appeared briefly, within 50ms of a second segment from the same array (containing 12 separate dots), subjects’ visual systems would combine the percepts of these segments, enabling subjects to accurately locate the missing (25th) dot. However: if the second segment appeared 100ms after the first subjects would perform poorly – since this would now exceed the timescale of visible persistence, they would rely on working memory to represent the first segment and would (consequently) only keep track of (roughly) four dots from it (Di Lollo, 1980; Eriksen & Collins, 1967; Loftus & Irwin, 1998).
Amazingly, Brockmole et al. showed that, in spite of this, when the second segment appeared 1,500-5,000ms after the first, and subjects were told to try and “imagine the dots still being present after they had disappeared” (p.317), they would reliably identify the missing dot (almost as reliably as if all 24 dots were presented at once). This is significant for while 1,500ms is considerably longer than the 50ms the first segment would have visibly persisted for, it is independently predicted to be the length of time it takes subjects to construct a mental image (for discussion, see Kosslyn et al., 2006). So, according to the authors of the study, this provides reason to think a mental image was being constructed in light of subjects’ beliefs about their task, and that it was then combined with their percept of the second segment. As such, Block claims that these “results demonstrate a direct content-specific effect of cognition on perception” (2016, p.22). That is: a case of genuine cognitive penetration as we are interested in it (see also: Kosslyn et al., 2006, ch.3).

Of course, there remains room for doubt. For one thing, the mere fact that a percept can be combined with a mental image, fails to show that these are combined within perception – after all, we think about the things we see all the time. Nevertheless, the suggestion can be bolstered. For instance, in a forthcoming book, Block makes further appeal to a study by Pearson et al. (2008), examining the effects of mental imagery on binocular rivalry. Here, subjects had conflicting visual patterns presented to each eye. Under normal conditions this leads to a competitive interaction. That is: one pattern becomes dominant in vision for a given period of time, after which the dominant pattern (irresistibly) flips. Pearson et al. showed that merely imagining one of the two patterns facilitated its dominance during subsequent presentations (p.982). This revealed that mental imagery had a similar effect to cases of ‘weak perception’, facilitating dominance in the way an imagined pattern would have had it been presented to the subject at low (but not high) luminance levels (p.983). This is significant since the functional equivalence found here provides evidence that cognitively driven mental images interact with the same visual mechanisms as those involved in the determination of visual dominance and influence their operations in the same way as weak percepts derived from the retina. So, when considered alongside the Brockmole et al. study (mentioned above), these findings may provide grounds for thinking that mental images enter into genuine perceptual processing, modulating high-level visual functions, and thereby supporting the view that mental imagery can (or does) constitute cognitive penetration.

Now, let us suppose this is so, if only for the sake of argument. My concern is how this should bear on the hypothesis that perceptual systems are informationally encapsulated modules. Block (forthcoming), Prinz (2006) and others (Firestone & Scholl, 2016; Pylyshyn, 2002; Robbins, 2009) assume it would problematize – or even refute – the view. But, given the conclusions of §2, this
seems to be a mistake. Any cognitive penetration that is involved here is, in fact, consistent with the informational encapsulation of perceptual systems, and would even seem to indicate as much.

To illustrate, note that Brockmole et al. take their findings to vindicate the hypothesis that mental images and visual percepts are combined on a ‘visual buffer’ (p.329). The visual buffer is a hypothesised functional space onto which mid-level percepts and mental images can be projected and maintained (see Kosslyn, 1980; 1997). Once on the buffer these representations are treated equivalently by subsequent processing (e.g. high-level perceptual systems, like those involved in the determination of visual dominance).

Importantly, this is meant to explain how mental images and percepts can become integrated within the perceptual hierarchy, despite the processes involved in their production relying on independent bodies of information and requiring functional independence in their operations (see Kosslyn, 1980, ch.2). But: given §2.3, this renders the above findings consistent with the full-blown, informational encapsulation of each and every perceptual system. Since the visual buffer is effectively a joint between functionally independent systems in the visual hierarchy (housing the outputs of independently posited low/mid-level visual systems, before these get taken as inputs by independently posited high-level systems [see Marr, 1982, and, indeed, Prinz, 2012]) the integration of percepts and mental images here leaves untouched the idea that every perceptual system produces its outputs on the basis of a prescribed proprietary database which fails to include cognitive information. So, even if the integration of mental images and percepts on a visual buffer constitutes genuine cognitive penetration, it is not obvious why this threatens the informational encapsulation of perceptual systems.

Pearson et al.’s findings support this contention. Recall: the reason why they are said to indicate that mental imagery becomes integrated with perception in perception is that they purport to reveal the functional equivalence of mental imagery and cases of actual perception (specifically, cases of ‘weak perception’ in the determination of visual dominance). Pending a better explanation, this suggests that mental imagery interacts with the same perceptual mechanisms as genuine percepts when perceptual dominance is being determined. But note: when it is a fully formed percept (presented to the eye) that is interacting with these mechanisms, it would be uncontroversial to suppose that the percept (or some representational precursor) is merely acting as an input to the relevant system/process (after all, it is being derived from the retina). So, to the extent that mental images are processed in a functionally equivalent manner (i.e. the datum that supports their influencing perception proper) we should take this to be evidence that, in mental imagery, mental images are simply serving as inputs to the relevant mechanism(s) in much the same way (if they
were not, we would expect to find functional differences between the cases and the argument for cognitive penetration would not go through). But, pace Block and his sympathisers this leaves the findings of Pearson et al. consistent with the informational encapsulation of the system(s) involved, even if cognitive penetration is genuinely occurring. Since the mental images in question simply form inputs to the relevant systems determining dominance, and do not bear on the scope of their proprietary databases, these findings leave the encapsulation of the system(s) intact, in line with the proposal sketched in §2.

All of this seems to be more than a bare possibility. If a functional space, like the visual buffer, is the place where mental imagery becomes integrated into perceptual processing (as the aforementioned studies are meant to evince [see also: Kosslyn et al., 2006]), then we should want to know what it is that ensures that mental images are projected onto the visual buffer and not other locations in the perceptual hierarchy. Here, a natural and (by the Fodorian’s lights) independently motivated suggestion would be that it is the relevant systems’ informational encapsulation that ensures this. On this view, the relevant perceptual systems (those producing outputs on the buffer, and those taking the contents of the buffer as input) would process their inputs (derived from the retinae, visual buffer or wherever) on the basis of architecturally prescribed proprietary databases that prevent the admission of mental images (created on the fly) or the cognitive states that led to their creation. Consequently, there is simply nowhere else for a mental image to slot into, and thereby penetrate, perceptual processing, other than at the joints between modules (like the visual buffer). Pending a better explanation, these considerations then suggest that the above findings are not only consistent with the informational encapsulation of our perceptual systems, but actually indicative of this.

Opponents might respond by denying the existence of a visual buffer, as Kosslyn and others conceive of it. Instead, they might propose that mental images (or the cognitive states driving their production) enter into the bodies of information that guide our perceptual systems in the interpretation of their inputs, rather than just supplementing or modifying their inputs. But note: it is the hypothesis that there is a visual buffer, onto which mental images are projected, that the aforementioned studies (cited by my opponents) are designed to test and evince. Indeed, the argument for cognitive penetration that is based on these studies seems to rely on this. After all, in the Pearson et al. study that Block highlights, it is the functional equivalence of mental imagery and weak perception that indicates that mental imagery is modulating genuine perceptual processing. But, as we have just seen, this implies that mental imagery merely forms an input to the affected system(s), just as the visual buffer hypothesis predicts.
A more promising response may then involve finding independently motivated explanations for integration on the visual buffer that does not appeal to the informational encapsulation of the relevant perceptual systems. Quite what these explanations might be remains unclear, however. Opponents of informational encapsulation have provided many reasons why unencapsulated systems might often appear to be encapsulated, and uninfluenced by background knowledge of some variety, despite possessing access to this information. However, standard accounts of this sort seem irrelevant to the above cases. For instance: in response to the (apparent) judgement independence of the Muller-Lyer illusion (see §2.2), opponents of perceptual encapsulation have claimed that this simply reflects the fact that bottom-up sensory processing trumps top-down predictions within the perceptual hierarchy (Clark, 2015; Prinz, 2006), or that it merely reflects the sluggish speed at which cognitive influences occur (Shea, pers. comm.). The trouble is: neither of these proposals explicate the phenomena currently under discussion. In binocular rivalry and perceptual integration, the mental image has (allegedly) already entered into perceptual processing when it elicits its effect on visual processing, and an appeal to encapsulation accommodates (and, as we have seen, elucidates) this. Thus, the speed at which it enters into perceptual processing is irrelevant. Moreover, it is not the case that the cognitive effect is being trumped by bottom-up visual processing since (by hypothesis) the mental image – once formed – is being integrated into bottom-up visual processing and treated ‘equivalently’ to it. So, pending a better (and so-far lacking) explanation, Block and others’ alleged cases of cognitive penetration involving mental imagery should actually be seen to indicate that perceptual systems are informationally encapsulated modules, rather than refute this, even if cognitive penetration is genuinely occurring.

3.2 Critiquing the View

The preceding remarks recommend a surprising result. Cognitive penetration is often seen to present a deep problem for the modularity of perception, insofar as this requires the informational encapsulation of perceptual systems and perceptual computations. Given the above, this looks to be a mistake.

To illustrate, we considered the phenomenon of mental imagery. Mental imagery is often seen to provide a particularly plausible example of cognitive penetration. But any cognitive penetration it involves seems to be both consistent with, and even indicative of, a view on which perception is entirely constituted by informationally encapsulated modules, and on which every perceptual computation is encapsulated from central cognition. I say this because the penetration it (allegedly) involves simply concerns a cognitive (non-perceptual) process (the formation of a mental image) supplementing or modifying representations at non-arbitrary points in the perceptual hierarchy.
(specifically, on a visual buffer, which houses the outputs of independently posited intermediate-level perceptual systems [Kosslyn, 1980; Marr, 1982]). As we have seen, this is naturally accommodated, and even explained, by postulating that the relevant perceptual systems (both the low-mid level perceptual systems which produce outputs on the buffer, and the higher-level perceptual systems which take the contents of the buffer as input) are informationally encapsulated modules. Why? Because this would serve to explain why penetration simply occurs on the buffer, at the intersection between these systems, and not elsewhere in the hierarchy.

For what it’s worth, I am sympathetic to the thought that similar conclusions will apply elsewhere. For instance, various theorists have claimed that certain forms of top-down attention constitute forms of cognitive penetration, influencing processing within the perceptual hierarchy, as opposed to simply influencing inputs to the perceptual hierarchy as a whole (e.g. Block, 2016; Lupyan, 2016; Mole, 2015; Wu, 2014). This has been seen to present a further problem for the modularity of perceptual systems (see previous citations; but see also Quilty-Dunn, forthcoming a). Yet, since top-down perceptual attention seems unable to select anything more than a tightly constrained set of features, happenings and property-types, which (again) correspond to the outputs of independently posited mid-level perceptual systems (see Wolfe & Horrowitz, 2004; see also Green, forthcoming), proponents of perceptual modularity may (again) maintain that these effects simply result from cognitive influences on the joints between independently posited perceptual modules – perhaps, the very same joint as penetration mediated by mental imagery, further motivating the thought that this joint is special for lying between fully encapsulated perceptual modules.

In both cases, this seems like good news for proponents of perceptual modularity. But since my discussion constitutes a departure from standard ways of thinking about encapsulation and its commitments, I wish to close by highlighting two (non-exhaustive) ways in which critics might legitimately respond.

First: opponents might seek out evidence that cognitive penetration (also) occurs at arbitrary points in the perceptual hierarchy. By ‘arbitrary’ I mean points in the perceptual hierarchy that seem unlikely to correspond to the outputs of any interestingly domain specific process that might plausibly constitute a functionally specified module. This does not seem to be true of the cases we have considered since the cognitive penetration that is allegedly taking place there involved the elicitation of effects at a joint between independently posited modules – namely, on a visual buffer, at the joint between mid and high-level visual systems of the kind posited by Marr (1982), Jackendoff (1986), Prinz (2012) and many others. Plainly, this is not inevitable, however. To the extent that cognitive penetration could also be shown to elicit effects at arbitrary points in the perceptual
hierarchy (i.e. points that are unlikely to correspond to the joints between functionally independent sub-systems) this would undermine my hypothesis.

Second: critics could seek out cases of cognitive penetration which constitute perceptual computations (i.e. computations in perception that rely on access to central cognitive states). If successful, this would force a reassessment of the extent to which perceptual systems and perceptual computation is informationally encapsulated, and of the extent to which perceptual computations are solely the result of modular operations. But note, in the aforementioned cases of mental imagery this does not seem to be so. There, mental images may have been created in light of the organism’s cognitive states. And this may lead to their supplementing or modifying low-mid level sensory outputs. However, this supplementation/modification computation does not seem to be a perceptual computation. Rather, it seems to be a non-perceptual cognitive process (Gauker, 2011; Johnson-Laird, 2001; Kosslyn et al., 2006). In this way, it presents no challenge to the idea that every perceptual computation is carried out by an entirely encapsulated modular system and that perceptual processing remains an entirely modular affair.

4. Conclusion

In sum: Cognitive penetration is widely assumed to undermine the informational encapsulation of perceptual systems. However, I have argued that this is non-obvious. Cognitive penetration occurs if one’s cognitive states exert direct and semantically coherent effects on their perceptual processing (§2.1), while informational encapsulation obtains insofar as a system processes its inputs on the basis of an architecturally prescribed proprietary database that excludes information accessible to other psychological systems of the organism (§2.2). Despite an apparent tension, this allows that cognition might penetrate perception at the joints between modules in a fully modular perceptual hierarchy, despite leaving each and every perceptual system’s encapsulation from central cognition intact, and every perceptual computation modular (§2.3). Indeed, once this is recognised, particularly plausible cases of cognitive penetration, like those mediated by mental imagery (by some accounts, the most plausible cases of cognitive penetration) may actually seem to evince the encapsulation of perceptual systems, rather than refute or undermine this (§3).

In recognising this, we defuse a prominent objection to the modularity of human perceptual systems and find novel reason to endorse the view. Still, much work remains to be done. As noted, mental imagery is but one (alleged) form of cognitive penetration. Thus, researchers should consider other (alleged) cases and ask how (if at all) these considerations bear relevance. Somewhat independently, researchers may wish to examine how these findings bear on the possible encapsulation of other high-level systems, sometimes associated with holistic thought, like those
involved in flexible goal ascription (cf. Butterfill & Apperly, 2016) or cheater detection (cf. Van Lier et al., 2013). Additionally, they might consider more closely where the joints between (putative) modules in (e.g.) the visual system are (or should be) located, perhaps using this to characterise discrete stages of visual processing. In each case, our discussion illustrates that the crux of the modularity hypothesis is not whether cognitive penetration occurs; it is whether this threatens the existence of dedicated proprietary databases governing the relevant systems’ operations. This is a moral that is often missed, but which must be kept in mind when advocating for, or objecting to, the encapsulation of any given psychological system.

(8,917/10,843 words)

Works Cited:


4. – (forthcoming). *The Border Between Seeing and Thinking* (book manuscript)


