

*This is an abstract; Final paper to appear in Cognitive Effects on Perception: New Philosophical Perspectives, Ed. A. Raftopoulos and J. Ziembekeis*

Towards a consequentialist understanding of cognitive penetration

Dustin Stokes  
University of Toronto

The cognitive penetration of experience, if it occurs, is a phenomenon of significant philosophical and scientific importance. Whether the phenomenon occurs is debated. Much of this debate is unfortunately rooted in lack of consensus on just what the phenomenon is or would be. There is little to no agreement, on either side of the debate, regarding a definition or analysis of the target phenomenon. This has important consequences. Perhaps most importantly, and very simply, if cognitive penetration is to be empirically testable, then there must be some agreement on just what one is testing for (and thus how experiments should be designed and controlled). As it stands, empirical data are interpreted differently by different theorists, and by appeal to different criteria for what “counts” as a case of cognitive penetration. A variety of sceptical (that is, non-cognitive penetration) interpretations are invoked, and without an uncontroversial definition of cognitive penetration in hand, adjudication is difficult if not impossible.<sup>1</sup> Indeed, how could one claim that the sceptical interpretations are less plausible, and that an observed mental phenomenon is *best* explained as cognitive penetration, if the latter is not clearly defined or, better, not at least defined in a way that is agreeable to both parties of the debate. Unfortunately, a controversy partly predicated on cross-talk is not an unfamiliar scenario in philosophy (or science). But fortunately, in this case, there is a strategy for resolution and one that, hopefully, will encourage progress on an important (possible) aspect of the human mind.

---

<sup>1</sup> These interpretations, described by various theorists, include the memory interpretation, the attention (shift) interpretation, and the judgement interpretation. See Macpherson 2011; Stokes, forthcoming; and for examples some of these interpretations in use, see Fodor 1988; Pylyshyn 1999.

This paper first analyzes some extant definitions of cognitive penetration (or, at least characterizations) offered by both sides of the cognitive penetration debate. Perhaps the first definition of the phenomenon comes from Zenon Pylyshyn (1999; though see also Pylyshyn 1980; 1984). His definition places emphasis on a semantic coherence criterion whereby there must be some identifiable semantic coherence between the penetrating cognitive state and the penetrated perception. This is a useful qualification, since it distinguishes cases of mere causal influence (e.g. where a belief about one's surroundings causes one to look around and, consequently, enjoy changes in perceptual experience) from a more narrow class of (possible) mental phenomena. However, this condition is too strong in a variety of ways: depending upon interpretation, it commits to an insufficiently neutral account of cognition, or it unnecessarily imposes normative constraints *or* operationalist constraints on the target phenomenon. Fiona Macpherson's more recent proposal is something of an advance, since it explicitly attempts to define the phenomenon in a way that rules out alternative interpretations (Macpherson 2011). But to the degree that her definition also commits to Pylyshyn's semantic coherence criterion, her definition is too strong. One might avoid these challenges by resisting commitment to the semantic coherence criterion, as some recent proposals have done (Siegel 2011; Stokes, forthcoming). These definitions, however, appear too weak. In particular, the sufficiency of the relevant conditions in these definitions are open to counterexample. The familiar lesson here is that defining cognitive penetration (perhaps like any definiendum of substantial interest) is no easy task.

Research is further confounded by a variety of additional factors and considerations. First, how direct must cognitive penetration be? That is, must the background cognitive state affect perceptual experience without any mediating event? Or, as some have argued, may cognitive penetration involve a cognitive influence on perception mediated by some

additional mental state, for example, mental imagery (Macpherson 2011)? A related question concerns attention as a possible mechanism of penetration. It seems clear, as sceptics like Fodor have well-illustrated, that overtly shifting one's attentional gaze so as to effect a changed perceptual experience is not cognitive penetration in any interesting sense. But this sceptical interpretation does not apply to all possibly relevant cases. One might wonder about scenarios where a belief or desire causally influences (in a top-down way, as it is sometimes put) perceptual or pre-perceptual selection mechanisms, which in turn causally influence perceptual experience; would cases like these count as cognitive penetration? A third question concerns the candidate penetrators of perception. The states most discussed in the literature are doxastic states. But some have argued that non-doxastic states like desire or value may non-trivially influence perceptual experience (Stokes, forthcoming). Does this count as *cognitive* penetration? Finally, one may wonder about the relevant time-scale. Are genuine cases of cognitive penetration only *synchronic*: where a cognitive state influences perceptual experience (relatively) immediately? Or, as some have suggested, is *diachronic* cognitive penetration possible where, for example, learning affects perceptual experience over time (Churchland 1988; see also Siegel 2006 for a related discussion)? None of these questions are easily answered, most especially if we lack a suitable definition of cognitive penetration.

The diagnosis here is simple: attempts to define or characterize the phenomenon have erred towards defining cognitive penetration as such, and in a way that has lost sight of the supposed consequences of the (possible) phenomenon. That is to say, there are reasons that philosophers and cognitive scientists began discussing the possibility of something like cognitive penetration, and those reasons all concern the consequences that a phenomenon

would have *if* that phenomenon occurred (and, perhaps, with some frequency). There are three central consequences of interest.

The first was the subject of critics of empiricists models of science in the middle of the 20<sup>th</sup> century. A number of philosophers challenged the traditional assumption that scientific observation is theory-neutral in a way that would support rational theory choice (Hanson 1958, 1969; Kuhn 1962; Feyerabend 1962). At least one version of this worry understands the relevant observation as, simply, perceptual observation. The corresponding worry is epistemic: if perception is laden with theory (in particular the theory or theories being tested), then it will not provide a means for rationally choosing or adjudicating between scientific theories.

The second consequence is a general epistemic one. Perceptual experience, on one very compelling picture, provides us with knowledge about the world. But this epistemic role for perception is threatened if there are circumstances where background cognitive states influence perceptual experience that, in turn, influence belief formation. The most obviously pernicious cases are ones that involve a rough causal schema of the form:

$$\textit{Belief that } P \rightarrow \textit{Experience that } P \rightarrow \textit{Belief that } P$$

Here the causal history of the consequent belief involves a circularity that, plausibly, undermines reason for that very belief (Siegel 2011, forthcoming; Lyons, forthcoming).

The general epistemic worry is in fact of central relevance to the third possible consequence of cognitive penetration, namely, architectures of the mind. Modularity theorists (at least of Fodorian strength) claim that perceptual systems are *informationally encapsulated*. A core motivation (perhaps the *primary* motivation) for this claim is that the processing of such systems would be immune to error introduced by the broader cognitive system. In other words, since computations performed by these modules are supposed to be

insensitive to what the organism knows, expects, or wants, the resulting perceptual representations more reliably inform the organism about its environment (Fodor 1983: 68-70; see also Pylyshyn 1980). If, as it turns out, cognitive penetration (of perceptual systems) occurs then this alleged feature of modules, even if epistemically desirable, is not actual.

Now, setting aside various complications, these general issues initially motivated theorists to ask whether cognitive penetration occurs. In other words, the fundamental question(s) asked by theorists was whether some phenomenon or family of phenomena occurred that resulted in the above consequences. This *is* something about which most if not all parties of the debate agree. The thesis of this paper is that an analysis of cognitive penetration should be constrained accordingly. An analysis (even if not a full-blown *essential definition*) of cognitive penetrability will be successful just in case and to the degree that it (a) describes a phenomenon (or class of phenomena) that has implications for theory-ladenness and rational theory choice, the knowledge-providing role of perception, or the modularity of mind (or better: some combination thereof); and (b) describes a phenomenon (or class of phenomena) that is not aptly interpreted in any of the sceptical ways invoked in current literature. These constraints are two ways of getting at the same thing: the alternative interpretations are effective *because* they preclude the consequences of interest. But making both constraints explicit may be useful, since constraint (a) describes what cognitive penetration should be or imply, and constraint (b), what it should not be or imply.

The virtue of this consequentialist constraint on analyses of cognitive penetration is this: once such an analysis is in hand, empirical studies can be devised and executed accordingly. Testing for cognitive penetration becomes testing for a phenomenon that bears the relevant consequences for the epistemology and architecture of mind. This (re)identifies the phenomenon of interest as one common to both sides of the debate. Moreover, it would

provide some metric for answering the various “what counts as cognitive penetration?” questions described above. None of this, of course, settles the debate. But it does identify a conceptual strategy for refining theoretical and empirical research and, with some luck, making new progress on an important set of questions about the human mind.

#### Works cited

Churchland, P. M. ‘Perceptual Plasticity and Theoretical Neutrality: A Reply to Jerry Fodor.’ *Philosophy of Science* 55 (1988): 167-187.

Feyerabend, P. ‘Explanation, reduction and empiricism.’ *Minnesota studies in philosophy of science, Vol. 3* Ed. H. Feigl & G. Maxwell. Minneapolis, MN: University of Minnesota Press, 1962. 28-97.

Fodor, J. *Modularity of Mind*. Cambridge, MA: MIT Press, 1983.

\_\_\_\_\_. ‘A reply to Churchland’s ‘Perceptual plasticity and theoretical neutrality.’ *Philosophy of Science* 55 (1988): 188-198.

Hanson, N. R. *Patterns of Discovery*. Cambridge: Cambridge University Press, 1958.

\_\_\_\_\_. *Perception and Discovery: An Introduction to Scientific Inquiry*. Freeman and Cooper, 1969.

Lyons, J. ‘Circularity, Reliability, and Cognitive Penetrability of Perception.’ *Philosophical Issues* (forthcoming).

Macpherson, F. ‘Cognitive Penetration of Colour Experience: Rethinking the Issue in Light of an Indirect Mechanism.’ *Philosophy and Phenomenological Research* (2011).

Pylyshyn, Z. ‘Computation and Cognition: Issues in the Foundations of Cognitive Science.’ *Behavioral and Brain Sciences* 3 (1980): 111-32.

\_\_\_\_\_. *Computation and Cognition*. Cambridge, MA: MIT Press, 1984.

\_\_\_\_\_. ‘Is vision continuous with cognition? The case for cognitive impenetrability of visual perception.’ *Behavioral and Brain Sciences* 22 (1999): 341-365.

Siegel, S. ‘Which Properties are Represented in Perception?’ *Perceptual Experience*. Ed. T. Szabo Gendler and J. Hawthorne. Oxford: Oxford University Press, 2006. 481–503.

\_\_\_\_\_. ‘Cognitive Penetrability and Perceptual Justification.’ *Nous* (2011).

\_\_\_\_\_. ‘The Epistemic Impact of the Etiology of Belief.’ *Philosophical Studies* (forthcoming).

Stokes, D. ‘Perceiving and Desiring: A new look at the cognitive penetrability of experience.’ *Philosophical Studies* (forthcoming).