Cognitive penetrability of perception

Dustin Stokes
University of Utah
dustin.stokes@utah.edu

Abstract
Perception is typically distinguished from cognition. For example, seeing is importantly different from believing. And while what one sees clearly influences what one thinks, it is debateable whether what one believes and otherwise thinks can influence, in some direct and non-trivial way, what one sees. The latter possible relation is the cognitive penetration of perception. Cognitive penetration, if it occurs, has implications for philosophy of science, epistemology, philosophy of mind, and cognitive science. This paper offers an analysis of the phenomenon, its theoretical consequences, and a variety of experimental results and possible interpretations of them. The paper concludes by proposing some constraints for analyses and definitions of cognitive penetrability.

The dominant tradition in philosophy and psychology is to distinguish human perceptual experience from human cognition. Like many distinctions, this one is not readily made in an uncontroversial way. One method of distinction is to enumerate paradigmatic examples of each category. Such examples of perceiving include seeing, hearing, and touching. Examples of cognitive states or processes include believing, intending, and reasoning. Of course, we still want to know the criterion or criteria by which these mental types are put on one list rather than the other. There are at least two non-exclusive means by which this might be done: function and phenomenology.

Consider vision. It functions to provide an agent with information about the colour, shape, and location features of the objects of her immediate environment. Thus we say that when one suffers a visual hallucination, vision has malfunctioned. Beliefs, by contrast, do not perform a function that is obligated to one’s immediate environment. Although one can and does form beliefs about one’s current surroundings, one also forms and maintains beliefs about the past and the future, and about locations removed from what is perceptually

---

1 Thank you to Vince Bergeron, Fiona Macpherson, and Wayne Wu for discussion of some of these issues. Special thanks to Susanna Siegel for reading and commenting on an earlier draft. And thanks finally to an anonymous referee for this journal.
available. For instance, I can’t presently see Vancouver, because I am in Toronto. But I presently have, and can reflect upon, many beliefs about Vancouver. Furthermore, one forms beliefs about things that are not perceptible at all, for example, about angels or Platonic Universals or the Gross Domestic Product of Sweden. This has a relevant physiological basis: belief states do not depend in any direct way upon current activity in one’s sensory organs. One cannot see when one is completely blindfolded, but one can still have and consider one’s beliefs. Add to this the observation that belief states, by contrast with visual experiences, figure directly into one’s deliberative decision making. In making a decision on whether to, say, buy a new home, one considers the facts, as we say. And this is another way to say that one reviews what one believes to be the facts, and then makes a decision guided by those beliefs. All of this is to characterize important functional differences between vision and belief, two paradigmatic examples of, respectively, perception and cognition.

Alternatively or additionally, one might distinguish perception from cognition by appeal to phenomenology. There is something that it is like to have perceptual experiences (Nagel 1974). Visual, auditory and other sensory experiences are characterized by a qualitative phenomenology. It feels a distinctive way, from the first-person perspective of the perceiver, to taste a bleu cheese or hear wind blowing through a pine tree. Plausibly, states like belief lack a distinctive phenomenology. Many of our beliefs will not feel any way, from the first person perspective. You may hold beliefs with the following contents: ‘2 is the square root of 4’, ‘270 electoral college votes are required to win the US Presidential election’, ‘St. John’s is the oldest city in Canada’. These beliefs may be more or less important to you, and will play distinctive roles in your reasoning, but it does not feel any way to you, qualitatively, to hold or entertain them. There is nothing that it is like to believe
that ‘St. John’s is the oldest city in Canada.’ By contrast, seeing the city of St. John’s will be rich in qualitative detail; it is characterized by a phenomenology that the correlative beliefs lack.\(^2\)

Although the above is not intended as an uncontroversial or exhaustive analysis of the perception/cognition distinction, it should provide the reader with a working idea of the ways that typical examples from each mental category are distinguished. We can then ask about relations between members of the two roughly distinguished categories. Tradition has it that perception influences cognition: what we see, hear, touch and otherwise perceive influences what we believe, intend, desire, and so on. But this causal influence is supposed not to go the opposite way: you have qualitative perceptual experience of the world in a way that is (generally) independent of what you know, believe, intend, and so on.

This last point needs qualification. Cognitive states often do influence how one perceives the world. For example I want some chocolate and I believe that there is chocolate in the cupboard. This combination of belief and desire motivates me to go to the cupboard and get some chocolate. As a result of this simple casual chain of events, I have an enjoyable perceptual experience of chocolate (in fact several: I taste, smell, touch, and see chocolate). Tradition does not resist this explanation. But it does resist a more direct cognition-to-perception causal influence. The way the chocolate looks or tastes to me—these appearance properties—will be the same no matter what I believe about chocolate or about this particular bar of chocolate, and no matter any of my other thoughts. And this is supposed to be true across perceivers with suitably similar sensory systems: the chocolate tastes and looks the same to an expert chocolatier as it does to me, even granting that the chocolatier will know a great deal more about the chocolate and accordingly make richer and better

\(^2\) However, some recent theorists have argued that beliefs, like all other intentional mental states, have a phenomenal character. See Horgan and Tienson 2002; Loar 2003; Strawson 1994.
judgments about it. The perceptual experiences, then, are supposed to be causally sensitive to the stimulus—the chocolate—and the operation of our sensory organs, but not to the perceiver’s antecedent thoughts about chocolate.

Tradition thus takes perception to be *cognitively impenetrable*. Cognitive states do not directly affect the way we see, hear, taste and otherwise perceive the world. A *cognitive penetrability thesis* simply denies this traditional view: perception is, sometimes, penetrated by cognition. This paper aims to further clarify cognitive penetrability theses, identify their importance to philosophy and cognitive science, and analyze relevant empirical data.

I. Further characterizing cognitive penetration

Cognitive penetration is supposed to be more interesting than the entirely common scenario described in the chocolate example above. (It may be more interesting *because* it has greater theoretical consequences, as will be discussed in sections II and IV below.) One thing clearly missing in this case is a more direct causal connection, since here cognitive states cause a set of actions (going to the cupboard, opening the cupboard, and so on) that then cause the relevant perceptual experience. Current definitions in the literature attempt to clarify the nature of the required directness.

One way to begin to narrow the focus to the relevant phenomenon is to consider characterizations of the opposing cognitive *impenetrability* thesis. As Fiona Macpherson describes it, perception is cognitively impenetrable if it is not possible for any two perceivers (or for the same perceiver at different times) to have experiences with distinct content or character when one holds fixed the object or event of perception, the perceptual conditions (e.g. lighting), the spatial attention of the subject, and the conditions of the sensory organ(s) (see Macpherson 2012). The virtue of this characterization is that it identifies and sets to one
side various factors that should not contribute to a genuine case of cognitive penetration: differences in distal and proximal stimulus, differences in attention, and differences in the sensory organ(s). More will be said on this below, but if one or more of these factors are present, then such factors would plausibly explain the difference in experience (rather than a cognitive difference being the explanation). So if A sees x as F, and B sees x as G, but the visual organs of A and B are functioning differently, or A and B are focusing their attention on different parts of x, then these differences in sensory organ or attentional focus best explain the perceptual differences between A and B. Inverting this characterization: if in holding those same factors fixed, it is possible for two perceivers to have experiences with distinct content or character (e.g. seeing an object as differently coloured), then cognitive penetration is possible.

Zenon Pylyshyn, who coined the term ‘cognitive penetration’, suggests that cognitive penetration must be a “semantically coherent” relation. This criterion can be interpreted in at least two ways. First, though, it is instructive to consider Pylyshyn’s original motivations for introducing a concept of cognitive penetration.

“Functions are said to be cognitively impenetrable if they cannot be influenced by such purely cognitive factors as goals, beliefs, inferences, tacit knowledge, and so on. Such a criterion makes it possible to empirically separate the fixed capacities of mind (called its ‘functional architecture’) from the particular representations and algorithms used on specific occasions” (Pylyshyn 1980: 111).

Thus, initially, Pylyshyn took cognitive penetrability to distinguish cognitive phenomena (understood computationally) from cognitively impenetrable functional architecture (understood biologically). However, this criterion, as critics point out, is too weak. It would bear the result that very little is cognitively impenetrable in the specified way: heart rate, digestion, and galvanic skin response are all, by this criterion, cognitively penetrable, and thus cognitive
capacities (Haugeland 1980; Kyburg 1980; Rey 1980; Smith 1980). This encouraged Pylyshyn to make an amendment. “[P]rocesses carried out in the functional architecture are processes whose behavior requires no explanation in terms of semantic regularities—that is, in terms of rules and representations. That position...provides a basis for a criterion I call ‘cognitive penetrability’” (Pylyshyn 1984: 131). Thus cognitively penetrable systems are influenced by systems the explanation of which requires terms of rules and representation, and explanation of that influence will also require terms of rules and representations.

This brings us, finally, to Pylyshyn’s most recent characterization of cognitive penetration. “[I]f a system is cognitively penetrable 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” (Pylyshyn 1999: 343). On one interpretation, the idea here is that the belief, desire, or other background cognitive state must stand in an inference-supporting relation to the experience or perceptual process.

Suppose I am on a deserted island and I believe that the (unripe) banana in my hand is yellow. And suppose that I see the banana as yellow, and because of my belief that it is yellow. In this case, the perceptual experience has (roughly) the same content as the background belief and this semantic coherence allows for inference. This criterion ensures that cognitive penetration is not a mere causal connection between cognitive states and perception. One may worry, however, about the strength of the semantic criterion.  

---

3 A moment’s reflection should allow the reader to identify why this is so. Consider: Suppose I believe that there is a monster under my bed. This belief may cause my heart rate to increase. Or it may cause me to suffer indigestion. Therefore heart rate and digestion are both influenced by belief. Therefore (by Pylyshyn’s criterion) heart rate and digestion are cognitively penetrable and, so, cognitive. This constitutes a reductio of Pylyshyn’s initial criterion.

4 To anticipate a reply, one might say “Since ‘cognitive penetration’ is a term introduced by Pylyshyn, he can stipulate whatever criteria he likes. His term of art, his definition. End of story.” In one sense this is exactly right: Pylyshyn has introduced and defined a term and then simply applied it to various mental phenomena. So if he wants to use his terminology this way, he is entitled to do so. However, as will become clearer in section II, there is now an active debate in the cognitive sciences—that Pylyshyn is party to—concerning whether there
Consider another debate in the philosophy of perception. An epistemic argument for the claim that perceptual experience is conceptual goes as follows. If perceptual experience provides reason for belief (and thus knowledge), then experience must be structured by concepts. Since experience does provide reason for belief, experience must be structured by concepts (McDowell 1994). Never mind the success of this argument; its operative conditional premise is motivated by the thought that in order to stand in a rational reason-conferring relation with belief, experience must stand in inferential relations with belief. And to do this, experience must be similarly structured. Thus to appropriately “hook up” with belief, the content of perceptual experience must, like belief contents, have a conceptual structure. So my belief that “The tomato is red” is justified by my experience of the object before me only if the representational content of that experience is somehow structured by the concepts ‘TOMATO’ and ‘RED’.

In this argument the need for an inferential relation is driven by a supposed normative relation running, asymmetrically, from perception to belief. Cognitive penetration runs the other direction, but if it is a normative relation, then the same principle could be used to motivate Pylyshyn’s semantic criterion. But cognitive penetration is not a normative relation. Indeed perhaps quite the contrary as discussed below, cognitive penetration may be epistemically pernicious. So Pylyshyn’s semantic criterion requires some alternative argument.

---

5 This is a substantial literature. See Crane 1992; Brewer 1999; Heck 2000; Byrne 2005. And see both Bermúdez and Cahen 2011 and Siegel 2010 for useful summaries of this and related debates.

6 Pylyshyn does sometimes talk about the relation as a rational one (or quasi-rational one), but there is little reason to think that ‘rational’ here is being used in a normative sense. And if it is, there is little reason to think that such a description is accurate. The question is whether cognition directly influences perception in some non-trivial way. There is no motivation for the claim that this influence could only be a reason-guided, rational one. We might prefer this, epistemically speaking, but there is little empirical reason to think that we are such lucky organisms.
Or perhaps Pylyshyn’s semantic criterion requires an alternative strength. The condition might instead be interpreted in a way that emphasizes representational content without emphasis on inferential relations between cognitive and perceptual states. In at least one place Pylyshyn (1999) does suggest this interpretation, “This is the essence of what we mean by cognitive penetration: it is an influence that is coherent…when the meaning of the representation is taken into account” (365, note 3). This relation needn’t be rational or logical in the sense that it preserves truth. Instead, the content of the belief (or other prior cognitive state)—what it is about—simply must have a non-arbitrary effect on perception. So this is a kind of operationalist reading of the semantic criterion. Cognitive penetration requires that we could (perhaps in fairly idealized conditions) identify how the representational content of the background cognition has affected the content of perception.

This version of the semantic criterion may be more plausible, but its motivation may still be doubted. Representational intelligibility is not a theory-neutral condition on cognitive penetration. Instead, it betrays Pylyshyn’s more fundamental (and earlier) motivation. The goal for Pylyshyn (1980; 1984) was to delineate the explananda for cognitive science along computationalist lines. According to the latter understanding of cognitive science, the discipline is committed to a doctrine: cognition is computation. And Pylyshyn aimed to capture this by distinguishing cognitive functions, by appeal to his semantic criterion, from hard-wired biological architecture. The consequence is that perception (or at least some of it), Pylyshyn argues, is not among the cognitive phenomena (since it fails to meet that semantic criterion for cognitive penetration). The trouble is that the computationalist theory of mind (and correlative understanding of cognitive science, as such) is a controversial one. And so we are being offered a theory-dependent characterization of a phenomenon that is, plausibly, theory-neutral. In short, not only computationalists are interested in cognitive
penetration. So the operationalist semantic criterion seems insufficiently motivated as a constraint on the metaphysics (or for that matter, physical facts) of the mind.\(^7\)

This lands us back where we started: cognitive penetration is a causal relation, but this is just a start. Pylyshyn’s semantic criterion attempts to supplement the causal relation, but the criterion is questionable. What is needed is a further condition (or conditions) that qualify the causal relation appropriately but without unnecessary appeal to normative or operationalist considerations. Generalizing from Stokes (2012), a possible definition goes as follows.

\[
(CP) \text{ A perceptual experience } E \text{ is cognitively penetrated if and only if (1) } E \text{ is causally dependent upon some cognitive state } C \text{ and (2) the causal link between } E \text{ and } C \text{ is internal and mental.}
\]

This definition makes clear what is ambiguous in other accounts, namely, that the (penetrated) relatum is perception at the level of experience. And the causal relation between experience and cognition is one of dependence. One simple way to understand this is counterfactually: if C did not occur (antecedent to E), then E would not occur. Thus the phenomenal character of one’s visual or auditory or other perceptual experience depends non-trivially upon a background belief, desire, or other cognitive state.\(^8\) This leaves

\(^7\) One way to see this, is to consider some of the important consequences of cognitive penetrability: modularity of mind and the knowledge-providing role of perception. These consequences are discussed in section II below, but for now the point is just this: cognitive penetration, if actual, will bear consequences for modular architectures of mind and for the supposed epistemic roles of perception, no matter whether it is a phenomenon that we can adequately test. So testability (by representational intelligibility) cannot constrain the defined nature of the target phenomenon.

\(^8\) Note here that desires are included among cognitive states. Some maintain that desires are not cognitive states, but instead emotive or conative. This concern can be set to one side. First, as one finds in the work of cognitive impenetrability theorists like Pylyshyn (see quotations above on p. 5 and 6; see also Fodor 1983: 68, 73), states like goals and desires are included among those that, if they directly affected perception, would render perception cognitively penetrable. And second, Pylyshyn and others maintain this claim for good reason: if desires directly affect perception, there will be important consequences for science, epistemology, and theories of mind (no matter whether desires are properly called ‘cognitive’).
significant wiggle room—for instance, C may cause intermediary cognitive states, which then cause E. At the same time, (2) further qualifies this causal link in order to rule out trivial cases and overt attention-shift cases. The causal chain must run from C to E without deviating from a mental series of events internal to the perceiving subject. Consider: I want to see my dog and I know that he is in the corner of the room. So I focus my attention to the corner of the room and consequently see my dog. Here the causal link from my cognitive states to my visual experience is mediated by an overt act of attention. By (CP), this is not an instance of cognitive penetration. In fact, a definition like this has the advantage of ruling out a variety of alternative interpretations that critics use to reject alleged cases of cognitive penetration. So evidence for experiences that meet CP would plausibly be evidence that both sides of the debate would have to accept.

As might be expected, this definition may not be immune to counterexample. One worry concerns the sufficiency of the two specified conditions. For instance, an example of Macpherson’s might be adjusted and invoked here. Suppose I suffer extreme exam anxiety and I believe that I am about to take an exam. This belief causes, internally, another mental state, namely the pain that accompanies a migraine. This pain further causes, again internally, a series of visual experiences where everything appears in a reddish hue. On the face of it, this scenario satisfies CP. But one may worry that this is not an instance of the relevant cognitive influence, since the causal chain from cognitive state to perceptual experience takes a circuitous route—even if internal and mental. Macpherson’s view is that this kind of case re-emphasizes the need for something like Pylyshyn’s semantic criterion: since the red hue experience in no way semantically coheres with the belief about the exam, it is not a case of cognitive penetration (Macpherson 2012: 26).

These alternatives are discussed in section III below.
The moral here is one familiar to philosophy: defining a phenomenon, however intuitive it may seem upon initial gloss, is no easy task. Nonetheless the above should suffice to provide the reader with a working understanding of cognitive penetration. And, as suggested below, considering the supposed consequences of the phenomenon and some relevant empirical studies sheds additional light.

II. Why cognitive penetrability matters

An alternative method for clarifying the phenomenon is to turn away from definitions and instead to reasons for which the phenomenon—even if imprecisely characterized—is supposed to be of importance to philosophy and cognitive science. There are at least three supposed consequences of the phenomenon: implications for theory ladenness and rational theory choice, the knowledge-providing role of perception, and mental architecture.

Beginning in the 1950s, philosophers of science challenged a traditional empiricist tenet regarding how scientific theories are chosen (Hanson 1958, 1969; Kuhn 1962; Feyerabend 1962). Assuming that, say, two incompatible scientific theories are equally sound—that is, both theories are internally coherent, offer clear hypotheses and predictions, are equally simple and parsimonious, and so on—then there must be some final court of appeal for adjudicating between them. The empiricist tradition suggests an obvious and intuitive solution: test both theories by comparing their respective hypotheses and predictions with perceptual observation of the world. And the theory that best comports with these observations is the one that we should, rationally, accept. Thus rational theory choice requires that perceptual observation achieves some significant objectivity.
The empiricist assumption identified and challenged by Hanson and others, is that perceptual observation (and, for that matter, judgements of and reports about perception) is neutral with respect to the theories being tested; it can be used safely and reliably as evidence for or against a theory. Hanson et. al. argued that this assumption was false: perception is laden with theory.\textsuperscript{10} This claim, if true, undermines the basic empiricist tenet: one cannot rationally adjudicate between theories on the basis of perceptual observation if these observations are already infected with beliefs and other mental states. So the cognitive penetration of perception threatens the scientific enterprise, where the penetrating cognitive states are doxastic commitments to one or another scientific theory.\textsuperscript{11}

This challenge generalizes to a second epistemic consequence. Another traditional view in western philosophy is that perception, at least in the majority of cases, can provide knowledge. When one has an ordinary visual experience—one’s eyes and visual cortex are functioning properly, the lighting is normal, there are no tricks, mirrors, or other illusion-inducing features of the environment—one typically forms beliefs on the basis of this experience. And as contemporary epistemologists put it, one’s perceptual experience epistemically justifies one’s belief. In the case that this belief is true—which, again, in the majority of cases it will be—one thereby comes away with some knowledge about the world.\textsuperscript{12} Some empiricists, moreover, take perceptual experience to provide a foundation for knowledge. Our knowledge about the world will, ultimately, be based on experience, with

\textsuperscript{10} In fact, these philosophers maintain a number of claims, varying in strength regarding theory-ladenness. For example, Hanson’s view was that perceptual experience itself is generally theory-laden. Kuhn worried about the theory-ladenness of perceptual judgement and observational reports, in addition to experience.
\textsuperscript{11} For a recent discussion, see Brewer and Lambert 2001. Their view is that scientific observation is theory-laden, but only when the subject of observation is ambiguous or somehow informationally degraded.
\textsuperscript{12} The reader will note that this way of framing the discussion begs a number of highly debated questions. To name three: questions about general skepticism; questions about internalist vs. externalist epistemic justification (for classic sources see: Alston 1989; Chisholm 1977; Dretske 1971; Armstrong 1973; Goldman 1979; for a recent collection see: Kornblith 2001); and finally the adequacy of the standard (JTB) account that takes knowledge to be some variant of justified true belief (Gettier 1963; and for a recent alternative to JTB, see Williamson 2000).
perception providing the terminus in the long chain of reasons that support one's putative knowledge about the world.

However, if perception can be infected with background beliefs and other cognitive states, then the supposed epistemic role of perception is threatened. Susanna Siegel provides an instructive example. “The challenge to perceptual justification posed by cognitive penetrability arises because it seems to introduce a circular structure to belief-formation…For instance, suppose Jill believes that Jack is angry at her, and this makes her experience his face as expressing anger. Now suppose she takes her cognitively penetrated experience at face value, as additional support for her belief that Jack is angry at him (just look at his face!). She seems to have moved in a circle, starting out with the penetrating belief, and ending up with the same belief, via having an experience” (Siegel 2011). This circular belief formation is epistemically pernicious. And the situation is even more pernicious if, as some have argued, desires, hopes, wishes and other non-doxastic states penetrate perception (Stokes 2012). The fact that I want some proposition P to be true—say, I want to be taller than you—provides no reason to have a belief that P. Wishful thinking may be common, and in some cases it may be practical, but it isn’t the sort of thinking that ever provides knowledge. I can’t know that I am taller than you on the basis of a desire with the same content. So, plausibly, if a perceptual experience is cognitively penetrated by desire or some other orectic mental state—importantly, such that the experience would not be had without that desire—then that experience cannot justify a belief or provide knowledge.¹³

¹³ Siegel (forthcoming) maintains that it is a substantive question whether certain (irrational) etiologies of perception, which mirror some epistemically problematic etiologies of belief, are similarly epistemically problematic. She argues that some etiologies of experience are rationally assessable, and take the same form as epistemically problematic belief etiologies. Among these are cognitively penetrated experiences, including those penetrated by desire. These experiences, by virtue of their “checkered past”, downgrade the justification-conferring role of the relevant experiences. See also Lyons (2011).
It is important here to identify some psychological and neurological phenomena whose relevance to these epistemic worries is controversial. In a debate with Jerry Fodor, Paul Churchland invokes phenomena where perceptual systems employ "theoretical assumptions" (Churchland 1979; 1988; Fodor 1984; 1988). Churchland asks us to consider phenomena like colour constancy—for example, where there are differences in illumination on two distinct points of a uniformly red tomato, one still sees the tomato as being red at both points—and amodal completion—where an occluded object, say a cat behind a picket fence, is visually perceived as a whole object in spite of some of its parts not being visible to the perceiver. The perceptual system, Churchland suggests, must make a number of assumptions about what is perceived, and this renders perception highly theoretical.14 Fodor, who maintains the cognitive impenetrability of perception, argues that these "assumptions", although clearly operative in perceptual processing, leave "perception neutral with respect to almost all theoretical disputes, [and so] couldn't ground any general argument for the unreliability of observation" (Fodor 1988:189). These features of the visual system are perhaps plastic such that they are shaped by encounters with one's environment. But these are not changes that human perceivers in any relevant sense learn, and have nothing to do with beliefs, theories, or other doxastic commitments that we may have. So Churchland's "assumptions" leave the epistemic normativity of perception untouched.

This raises a second issue: how does learning affect perception over time? The above discussion proceeds on the assumption that cognitive penetration, if it occurs and has the relevant epistemic consequences, occurs (relatively) synchronically: a cognitive state immediately influences one's present perceptual experience, and the putative theory-neutrality of perception is thereby undermined. Churchland, in the same debate, argues that

14 See Wu (forthcoming) for a discussion of spatial constancy and cognitive penetration, where he considers whether intentions (synchronically) affect perceptual constancy mechanisms.
diachronic penetration of perception may be equally relevant. To take one example, consider the way that perceivers adapt to the use of inverting lenses—lenses that, in short, turn one’s visible world upside down. On the standard story, the lenses are initially radically disorienting for subjects, impairing both judgement and eye-coordinated action. But after a learning period of consistently using and acting with the lenses, typically a week, subjects quickly recover to perform perceptual and motor tasks normally (Stratton 1897; Kottenhoff 1957; Taylor 1962). Churchland interprets this as penetration of perception across time: subjects learn new relations between movement and visual experience, re-formulate expectations accordingly and, eventually, perceive the world as they did before wearing the inverting lenses.¹⁵

A related debate concerns the admissible contents of experience. Recently, some philosophers have argued that perceptual experience represents not just low-level properties like colours and shapes (in the case of vision), but also high-level properties like natural kind, agential, and causal properties.¹⁶ One of Susanna Siegel’s arguments for high-level contents relies on cases where the phenomenology of experience allegedly changes as a result of the acquisition of recognitional capacities or beliefs. What it is like subjectively to look at Cyrillic text, as well as hear the language when spoken, will change for you after you learn the Russian language. The face of a friend as experienced will differ phenomenologically from

¹⁵ For this case to be relevant, the adaptation must be of one kind rather than another. Churchland’s assumption is that the phenomenal character of experience, after adaptation, returns to the right-side up character it had prior to the lenses. But there is another possibility: the phenomenology of experience with the goggles may stay the same—that is, upside down—but one learns, post-adaptation, to “deal with it” as we say. The descriptions by Stratton are, at best, ambiguous between these two possibilities (and in some places favour the second). See Hurley and Noe 2003, who argue for the former interpretation; and see Prinz 2006 and Schwitzgebel (manuscript) for skepticism regarding that interpretation. For a related discussion see McCauley and Henrich 2006.

when you first made the friend. These changes in phenomenology are common, depend on cognitive processes, and take place gradually over time (Siegel 2006).

The cognitive impenetrability theorist may resist these diachronic cases by noting that the alleged effects on perception result from a long-term learning or conditioning process, and thus lack the immediate cognition-to-perception relation necessary for cognitive penetration. Churchland’s response is that concerns about theory-ladenness are not concerns about instantaneous changes in belief (which then effect a change on perception). They are instead concerns about the observational effects of long-term and regularly evolving theoretical commitments. So if it is theory-ladenness that is (at least partly) at stake, then an insistence on synchronic cognitive penetration is unmotivated. More defensibly, Fodor insists that the changes to perception that occur in scenarios like the above involve a recalibration of the sensory system by the sensory system. There are good evolutionary reasons for this kind of plasticity, and none of it implies cognitive penetrability (Fodor 1988). It is difficult to determine whether inverting lens cases or cases involving change in recognitional capacity are instances of cognitive penetration properly called. But what matters most in this context of discussion is instead whether these phenomena imply the consequences for scientific theorizing and knowledge (discussed above).

In fact some have argued that cognitive penetration of perception is unnecessary to generate worries about theory-ladenness and the general epistemic role of perception. Even supposing that experience is impenetrable by higher-level cognitive states, the same concerns may arise if perceptual beliefs are immediately and robustly influenced by background cognitive states and theories. So, even if Tycho Brahe (who believes in an earth-centered universe) and Johannes Kepler (who believes in a sun-centered universe) see the

---

17 See Crutchfield (2011) for discussion of the relation between admissible contents of experience and cognitive penetration.
same thing as they watch the sunrise together—their visual experiences are, let’s suppose, phenomenologically identical—their background theoretical commitments may immediately influence the judgements made about these experiences. Brahe will judge and report seeing that the sun is moving vertically relative to a stationary horizon. Kepler will judge and report seeing that the horizon is moving vertically to expose a stationary sun (Hanson 1958: 6-8). In such a case, perceptual judgement is insufficiently theory-neutral to provide adjudication between the competing theories (again, even if phenomenal experience itself is theory-neutral). In a recent paper, Jack Lyons makes the same claim about the general epistemic worry, suggesting that perceptual beliefs are regularly penetrated by background cognitive states. The question then becomes, which kinds of influences are epistemically pernicious, and by what epistemic standards (Lyons, 2011; see also Churchland 1979; 1988; Siegel, forthcoming).

The third important consequence concerns implications for mental architecture, and this does require some cognitive effect on perception itself. According to some modularity theories of mind, perceptual systems are domain-specific, functionally independent systems. So there is a visual module, auditory module, olfactory module and so on (though the individuation of modules could be more fine-grained). These modules are, in their strong form, informationally encapsulated. Although a perceptual module \( m \) may exchange input and output with other systems in the organism, \( m \) cannot, in the course of its processing, compute over information available in other modules or systems. Encapsulated modules are thus supposed to be cognitively impenetrable: cognitive systems are among those systems relative to which perceptual modules are computationally isolated (Fodor 1983).

---

18 For a related empirical study, see Gunstone and White 1981.
modularity of mind is an empirical thesis, evidence for cognitive penetration of perception will count against the truth of the thesis.

Informational encapsulation of modules is often described at the level of computational processing or mechanism. Much of Fodor’s emphasis is on the computational mechanisms of perceptual “input systems”, and their lack of access to information available in “central cognitive processors” (Fodor 1983). And Pylyshyn’s main line of defence for the cognitive impenetrability of perception concerns early vision: a functionally defined component in visual processing that computes 3D shape descriptions of objects (Pylyshyn 1999). The rhetoric in this discussion often suggests that evidence for the cognitive penetration of perceptual experience bears no consequence for encapsulated modularity of perceptual processes. But there are at least two reasons to think that the opposite is true.

First, Pylyshyn and others are right to emphasize the fact that the cognitive penetration of some components of perceptual processing does not, by itself, imply the cognitive penetration of experience, since conscious experience may be the result of or be identified with some broader class of processing, and certain subsets of perceptual processing may not result in conscious experience at all.¹⁹ For the same reason, one cannot infer from the apparent penetration of experience to the penetration of any particular stage in perceptual computation. However, the cognitive penetration of experience implies the cognitive penetration of perceptual processing at some stage. Whether one takes experience to be identified with, constituted by, (metaphysically) determined by, or the output of perceptual processes, a difference in perceptual experience implies a difference in perceptual process. This implication holds no matter how one’s metaphysics of mind varies according

---

¹⁹ Some visual processing might, according to some theories, result only in sub-personal motor-guidance. See Milner and Goodale 1995; Clark 2001; Campbell 2002; Matthen 2005.
to these alternatives. So if experience is penetrated, then information processed (or outputted) by cognitive systems directly influences the processing of perceptual systems.

Second, it is important to recall some primary motivations for encapsulated modules. As Fodor puts it, the envisioned modular architecture would be evolutionarily advantageous, since it involves the “isolation of perceptual analysis from certain effects of background belief and set; and…this has implications for both the speed and objectivity of perceptual integration” (Fodor 1983: 43; emphasis added). Since an encapsulated perceptual module rigidly performs its function and with no interference from extraneous higher-level information—what the organism knows, expects, wants—it can rapidly provide accurate perceptual representations to the organism (Fodor 1983: 68-70; see also Pylyshyn 1980). It is this second epistemic point that is relevant here. A concern with the reliability or accuracy of perception is a concern with perceptual representation or experience, not merely processing. It makes sense to discuss the reliability of perception only in these terms: personal-level (mis)representations, not sub-personal computational mechanisms, lead or mislead the perceiver. One forms beliefs on the basis of what one sees, hears, and so on. So insofar as the modularity theorist is largely motivated by this alleged benefit of encapsulated perceptual modules, there must be a concern with perception at the level of experience.20

III. Evidence and interpretation21

---

20 For additional examples of strong modularity theories, see Fodor 1985 and Sperber 1996. For more on the relation between modularity and cognitive penetration, see Macpherson (2012); and Stokes and Bergeron (manuscript).

21 As should be expected, this section in no way provides an exhaustive review of relevant empirical literature, nor does it provide complete reviews of the studies discussed. Instead, it provides an overview of a variety of data—both old and new—that (a) are representative of some of the experimental methodologies that illuminate (even if they do not provide existence proofs for) the target phenomenon and (b) in many cases are illustrative of a variety of alternative (non-cognitive penetration) interpretations.
Evidence for cognitive penetration is contested. The goal here is to offer a brief survey of some of the debated evidence, and then to glean a number of skeptical interpretations that oppose a cognitive penetrability thesis.

Some have argued that top-down or “reentrant” pathways—neural connections that, evidently, enable signals from brain areas believed to process higher or conceptual information to more primitive areas—provide evidence for cognitive penetration (Churchland 1988, 1989; see also DiLollo et. al. 2000). While this remains an empirical possibility, it is far from sufficiently proved. As a number of authors have argued, this inference would require a relatively uncontroversial mapping from mental functions or states onto neural structures, and neuroscience is far from achieving this. So while the data on neural pathways may be suggestive, it underdetermines any conclusions about cognitive penetrability (Fodor 1988; Gilman 1991; Pylyshyn 1999; Raftopoulos 2001). By the same token, current neurological and neurocomputational data are insufficient to prove that perception is *im*penetrable.

Therefore, empirical evidence for (or against) cognitive penetration is best drawn from considerations and studies at the behavioural or psychological, rather than neurological, level. Much of this data, since it relies heavily on subjective reporting, suffers from the usual possible confounds and is open to a number of alternative interpretations.

Early defenders of cognitive penetrability relied heavily on evidence, largely anecdotal, concerning familiar illusions and ambiguous figures. For example, Hanson (1958) suggests that when one looks at the famous old woman/young woman reversible figure (see Figure 1), one can wilfully switch between two distinct visual experiences. And an early study showed that subjects primed with unambiguous photos of either an older woman or a

---

22 Kuhn 1962 and Churchland 1988 make similar appeals.
young woman were far more likely to see, respectively, the figure as an old woman or a young woman (Leeper 1935). While seeing the old woman is a distinct visual experience from seeing the young woman, in spite of the same stimulus (Boring 1930), the critic can resist these cases in one of two ways. The experiments in question show, one can argue, certain priming effects on perceptual receptivity, but this is not an effect of a belief or background theory. According to this *intra-perceptual interpretation*, the change in perceptual experience or processing is one effected by perception itself. Second, critics can respond to cases like this with an *attention-shift interpretation*. What one does when one “switches” between the old and young woman experience is to deliberately focus one’s attention on different points of the drawing which, in turn, results in distinct perceptual experiences. But this scenario is no different in kind from one where cognitive states like belief and desire motivate an action(s) which in turn results in changed experiences. I want the beer; so I drink the beer; then I taste the beer. If this is cognitive penetration then the concept is trivial, for the simple reason that we very frequently change our experiences by performing bodily actions or overtly shifting attention. And cognitive penetration is supposed to be a less ordinary phenomenon with, accordingly, special implications.\(^{23}\)

Hanson and others also appealed to the work of the New Look psychologists of the middle 20\(^{th}\) century. These psychologists, led largely by Jerome Bruner, held a view that perception and cognition were entirely continuous with one another; one’s perception is always framed by one’s background mental set—one’s beliefs, desires, expectations, conceptual categories, and so on. This universally quantified claim about human perception is easy prey to counterexample. Modularity theorists are fond of invoking cases where perception appears rigidly encapsulated. For example, although one may know that the lines

\(^{23}\) For an example of this reductio, see Fodor 1988: 191.
in the Muller-Lyer illusion are of the same length, one cannot see them as being the same length (see Figure 2).\textsuperscript{24} As a general psychological theory, then, New Look has been mostly abandoned. But it should be noted that many of the individual studies remain plausible evidence for an existentially quantified cognitive penetration thesis.\textsuperscript{25} And the opposing modularity thesis, as an empirical hypothesis, is an insufficient reason to reject any of the New Look data or more recent studies inspired by New Look. Therefore, evidence must be taken on a case-by-case basis, and inferences drawn must be abductive.

Balcetis and Dunning (2006) performed a series of studies aimed at testing whether desires influence perceptual experience.\textsuperscript{26} In some of these studies, subjects were briefly exposed to an ambiguous figure and then asked to categorize the figure. Depending upon how they categorized the figure, subjects would (and knew that they would) receive either a desirable or undesirable food. For example, a subject might be rewarded with a glass of orange juice if she categorized Figure 3 (see below) as a letter, or instead with an apparently disgusting concoction if she categorized the same figure as a number. In all of the relevant studies, Balcetis and Dunning's results strongly suggest that the food preferences bias how the ambiguous figure is seen. However, since categorization reports are made after the figures are displayed, critics may object that in these cases subjects are reporting their memories of the experience rather than the experience itself. Thus the memory interpretation

\textsuperscript{24} Additionally, a variety of empirical research suggests instances where perception is not relevantly affected by background cognitive states. See Carter and Schooler 1949; Klein, Schlesigner, and Meister 1951; Lysak and Gilchrist 1955; Tajfel and Wilkes 1963.

\textsuperscript{25} In addition to those studies discussed below, relevant studies in the New Look spirit include: Bruner and Postman 1948; Dukes and Bevan 1952; Bruner and Rodrigues 1953; Blum 1957. See Bruner 1957 for an overview.

\textsuperscript{26} For an earlier set of desire/food-related studies, and criticism, see Epstein 1961; Lazarus et. al. 1953; Saugstad 1966, 1967; Wolitzky 1967.
concludes that these data only evidence the influence of desire on a cognitive state like memory.\(^{27}\)

A rich area of recent study has focused on apparent cognitive effects on visual spatial perception. Subjects experiencing fear from atop a steep hill (e.g. when asked to imagine descending the hill while standing on a skateboard) more greatly overestimate the slant of the hill (both in verbal report and visual matching task) (Stefanucci and Proffitt 2008). Similarly, subjects who both report greater present fear and rate high on standard acrophobia measures more greatly overestimate vertical distance and size when viewing from a high place (Stefanucci and Proffitt 2009). Subjects who perform poorly on a sports task make correlative spatial judgements. For example, subjects who do poorly on a series of American football field goal kicks comparatively underestimate (after kicking) the width of the goal posts (Witt and Dorsch 2009). And objects that perform immediate goals (a bottle of water when thirsty) are judged closer than unwanted objects both by metric report and an action-based report (Balcetis and Dunning 2010).\(^{28}\) Whether these data support a cognitive penetrability thesis is debatable. Because these experiments involve the report of spatial judgements, critics may invoke an alternative judgement interpretation. This interpretation has it that experience across control and experimental subjects of the relevant stimuli is unaffected by background cognitive states, but the judgements made or beliefs formed about those

---

\(^{27}\) Keith Payne and colleagues have performed a number of important studies on racial prejudice and its apparent effect on perception. For example, American subjects more frequently mis-classify a hand tool as a gun when primed by black faces (Payne 2001). But because these experiments involve methods where subjects make reports only after the target image (a handgun or a hand tool) has been masked, here too the memory interpretation might be invoked by defenders of cognitive impenetrability.

\(^{28}\) Another possible interpretation of some of this data invokes the supposed distinction between motor-guiding vision—processed in the brain’s dorsal stream—and consciously experienced descriptive vision—processed in the ventral stream (Milner and Goodale 1995). For example, some of the experiments in Balcetis and Dunning 2010 required subjects to make distance estimates by tossing a small beanbag at the desired object. One might explain this effect as not one on conscious descriptive visual experience, but merely one where desire directly guides action via the dorsal stream. This would not be cognitive penetration as most theorists understand it.
experiences (or about the experienced objects) varies with the relevant background cognitive states. But this, just like the above alleged effects on memory, only supports a relatively uncontroversial phenomenon, one where cognitive states influence other cognitive states.

So, there are a variety of promising data for cognitive penetrability theorists but also a variety of alternative interpretations for defenders of cognitive impenetrability. The best current data—in the sense that it more plausibly evades these alternative interpretive strategies—result from online experimental methods that involve perceptual reports simultaneous with target perceptual experiences.

A classic such case comes from the New Look psychologists. Bruner and Goodman 1947 found that children significantly overestimate the size of coins relative to cardboard discs of analogous size. The desire for money evidently influences the perception of money. In a more recent study inspired by New Look psychology, adult perceivers estimate identically sized discs in a way that substantially varies depending upon whether the discs contain negatively valenced images versus neutral or positively valenced images. The researchers conclude that the subjects have and apply background values to the varying images, and this influences size perception of the discs (van Ulzen et. al. 2008). In another recent set of studies, researchers have found that conceptual categories affect performance on achromatic/colour tasks. Hansen et. al. (2006) found that when subjects are asked to adjust a naturally (yellow) coloured banana shape to appear achromatic, they over-compensate past the point of perfect grey, and adjust the shape into the opponent blue-hue range. This suggests that even once the shape is in fact grey, it still appears yellow-ish to the subjects. Levin and Banaji (2006) found that identically luminant greyscale faces with

---

29 See Stokes (2012) for discussion of both of these studies.
30 For possibly confounding results, see Olkonnen et. al. 2008, who found the effects to vary according to the “naturalness” of the target object. For example, 3D textured banana images were over-adjusted into the
features typical of the face of a white person versus those typical of a black person appear
differently, as subjects report a lighter match (by adjusting the luminance of, in some
conditions, a face and in other conditions a uniformly grey rectangle) for the typical white
face and a darker match for the typical black face. This effect occurs even when the target
faces are labelled as ‘WHITE’ or ‘BLACK’ but the features of the faces are identical (and
racially ambiguous).\footnote{See Macpherson (2012) for extended discussion of D elk and Filenbaum 1965 (an early study similar to the Hansen et al studies) and Levin and Banaji 2006 studies. Macpherson also argues for an indirect mechanism for cognitive penetration, whereby background cognitive states influence the phenomenal character of mental imagery which, in turn, influences (or is subjectively indistinguishable from) the phenomenal character of perceptual experience.}

The common factor important to all of these studies is that they involve tasks where
subjects are asked to inspect what they currently perceive, and report on the basis of that
experience. So subjects are not reporting on the basis of memory. Moreover, there seems to
be little room to appeal to overt acts of attention to explain the results: the actions required
of subjects for task performance, and their perceptual fields, are generally the same across
control and experimental groups. And the effects all seem tied to undeniably cognitive
factors, for example, to knowledge about natural and artefactual kinds and to racial
stereotypes. Finally, though it is possible that the reports are of judgements made about
one’s experience, to maintain the judgement interpretation one must maintain that, in these
online circumstances, subjects are making reports that deviate from their current experience.

For example, a subject in the Levin and Banaji 2006 study would on this
interpretation have veridical colour experiences of a racially labelled face and of the
adjustable report rectangle. Recall that the subject adjusts (in distinct trials) the patch to be

\documentclass[12pt]{book}
\usepackage{amsmath}
\usepackage{amssymb}
\usepackage{graphicx}
\usepackage{hyperref}
\usepackage{natbib}
\usepackage{url}
\begin{document}
\chapter{Introduction}

In this chapter, we will introduce the research questions and hypotheses that will be
addressed in the following chapters. Our primary goal is to understand the cognitive
mechanisms that underlie the perception of colour and shape in natural and artificial
contexts. We will begin by reviewing the literature on the perception of colour in
natural contexts and then move on to the perception of colour in artificial contexts.

\section{The Perception of Colour in Natural Contexts}

The perception of colour in natural contexts is a complex and multifaceted process
that involves both physical and psychological factors. Physical factors include the
wavelengths of light that reach the retina, the properties of the eye, and the
properties of the environment. Psychological factors include the cognitive
representations of objects and their interactions with the environment.

\section{The Perception of Colour in Artificial Contexts}

The perception of colour in artificial contexts is similarly complex and
multifaceted. Artificial contexts can be defined as environments that are
created by humans and are not part of the natural world. Examples of
artificial contexts include computer graphics, virtual reality, and
interactive multimedia.

In computer graphics, colour is used to create visual

differentiation. For example, in a game environment, different objects
may be distinguished by their colour to help the player

\section{Conclusion}

In conclusion, we have introduced the research questions and hypotheses that will be
addressed in the following chapters. Our primary goal is to understand the cognitive
mechanisms that underlie the perception of colour and shape in natural and artificial
contexts. We will begin by reviewing the literature on the perception of colour in
natural contexts and then move on to the perception of colour in artificial contexts.

\bibliographystyle{apalike}
\bibliography{references}
\end{document}
objectively lighter than a ‘WHITE’ labelled face and objectively darker than a ‘BLACK’ labelled face. Since the judgement interpretation resists any alleged effect on perception, this subject visually perceives, for example, the ‘WHITE’ labelled face accurately, namely, as of darker luminance than the grey rectangle (which, again, she herself adjusted). Further, this interpretation would also maintain that the subject would be consistently making judgements incompatible with, but on the basis of, those current experiences—submitting a report that the ‘WHITE’ labelled face is of the same luminance as the grey rectangle while, simultaneously, veridically perceiving them as being distinct in luminance. This may be less plausible than the opposing cognitive penetrability thesis. Perhaps perception is affected in these subtle ways. And this explains the subject’s reports: she sees the two objects as being of the same luminance, and so reports them as matching (even though, for example, the ‘WHITE’ labelled face is objectively darker than the adjusted grey rectangle). The inference to the best explanation, or so some have argued, is that experience is sometimes penetrated by cognition.

IV. Conclusion

None of the above studies provide conclusive evidence for the cognitive penetration of perception. This should be unsurprising. First, the target phenomenon is of substantial theoretical importance, with many theories at stake. Even if perception is theory-neutral, theoretical beliefs are not. Second, there is no single, uncontroversial definition of the target phenomenon. Third, the phenomenon as described is premised on a distinction between cognition and perception that is underdetermined by the neuroanatomical and behavioural evidence. As theorists of the mind, we may have very good reasons to make the cognition-perception distinction, but there is little reason to think that it is empirically discoverable in
some robust way. Therefore, the concluding suggestion is that cognitive penetrability should be defined by appeal to its supposed philosophical and scientific importance, and then tested accordingly.

The supposed consequences of the phenomenon (as outlined in section II) and the various alternative cognitive impenetrability interpretations (as outlined in section III) work together. The skeptical interpretations are effective, if they apply, because they block the implications to the putative consequences. So, for example, if a phenomenon (e.g. a subjective report that two objects are now the same colour) is best interpreted as resulting from an effect on a memory (with perception unaffected), then it does not imply the relevant epistemic concerns and threat to modularity of mind. All of this is instructive: the consequences and the skeptical interpretations can be used to construct a definition agreeable to both sides of the debate.

Therefore, future analyses of cognitive penetrability should be constrained in the following ways. A definition or analysis of cognitive penetrability will be successful just in case and to the degree that it (a) describes a phenomenon that has implications for theory-ladenness and rational theory choice, or the knowledge-providing role of perception, or the modularity of mind (or better: some combination thereof); and (b) describes a phenomenon that is not aptly interpreted in any of the skeptical ways outlined above. These constraints are two ways of getting at the same thing: again, the alternative interpretations are effective because they preclude the consequences. But making them both 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. Furthermore, these constraints are admittedly not rigid; for example, no commitment is made regarding the number of consequences that must be implied by a good definition, nor what the nature of the implication must be. But these
constraints do provide an informed way to better isolate a phenomenon that, if actual, all parties agree, would be important for how we theorize the mind.

In turn, once such a definition or 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 is an ideal setting for collaborative theorizing of the mind, where philosophy is needed to identify the consequences and analysis of cognitive penetration, and empirical science is needed to design and perform suitable experiments. This combined effort promises a better understanding of how thought influences experience of the world around us.

Works Cited and Additional Reading


Klein, G. S., H.J. Schlesinger, & D.E. Meister. ‘The effect of personal values on


Schwitzgebel, E. ‘The problem of known illusion and the problem of unreportable illusion.’ (unpublished manuscript).


—— and V. Bergeron. ‘Modular architectures and Informational encapsulation: A dilemma.’ (unpublished manuscript).


