This is a penultimate draft. Final and definitive form will be published in the *Australasian Journal of Philosophy*.

ATTENTION AND THE COGNITIVE PENETRABILITY OF PERCEPTION

Dustin Stokes
Department of Philosophy, University of Utah

Abstract
One sceptical rejoinder to those who claim that sensory perception is cognitively penetrable is to appeal to the involvement of spatial attention. While the sceptic is correct that some putative cases are accurately deflected in this way, the rejoinder oversimplifies the possible roles that attention might play in relevant contexts. This paper identifies alternative ways that selective attention might play a role in cognitive effects on perception. What emerges is a plausible and well-evidenced mental schema that describes attention-mediated cognitive penetration.

Keywords:
Cognitive penetrability of perception, Top-down effects on perception,
Attention, The cognition/perception distinction, Epistemology of perception, Modularity of mind

Researchers in philosophy and cognitive science debate whether cognitive or "higher-level" states like belief, desire, and intention influence, in some important way, sensory experience: whether perception is *cognitively penetrable*. Does the very look of an object, say a painting, vary from perceiver to perceiver in a way that depends upon those perceivers' background beliefs, desires, values? This paper offers a new framing of an important point of the debate, how attention may or may not mediate cognition and perception in a way compatible with cognitive penetration.

§1 offers brief clarification of the cognition/perception distinction and cognitive penetration. §2 identifies one common rejoinder to alleged cases of cognitive penetration, the "attention-shift interpretation", and identifies alternative ways that attention could mediate cognition and perception, and plausibly amount to cognitive penetration. The aim here is partly to shift the burden of proof to the sceptic of cognitive penetration, but also to make a case for attention-mediated cognitive penetration by appeal to some recent empirical research. §3 concludes with lessons that re-frame the relevant dialectical space. This analysis illuminates various features of attention and its relation to both cognition and phenomenal consciousness, and therefore should be of interest to a broad range of theorists of the mind.

1. Characterizing cognitive penetration

There is no uncontroversial way to define the terms 'perception' and 'cognition'. For present purposes, a standard pair of lists can be given, followed by a stipulated but orthodox characterization. Perception includes sensory experiences: seeing, hearing, tasting, smelling, touching. Cognition includes beliefs, intentions, goals, and perhaps desires and decision-making. Typically, perceptual experiences involve activity in one or more sensory organs, the presence of some relevant external stimuli, and involve rich first-personal phenomenology. Cognitive states require neither current sensory activity in any particular modality nor the presence of relevant external stimuli, and if they involve phenomenology at all, it is relatively impoverished.

Uncontroversial types of perceptual-cognitive interaction are familiar and pervasive. First, causal interaction runs from perception to cognition. What you see, hear, and otherwise sense, affects the cognitive states you form: your beliefs, your intentions for action, the decisions you make. The causal arrow also runs the other direction: your beliefs, intentions, desires, via the performance of action, cause changes in your sensory experiences.

The controversial possible interaction between perception and cognition is this. Can cognition affect perception in more direct or more important ways (for example, without the mediation of a series of actions)? Suppose that you and I are looking at the same object (say a painting) or the same event (say a possible hand ball foul in a football match). Can the first-personal, subjective qualities of our respective visual experiences differ because we hold different beliefs or values?

The dominant view in the last several decades of empirically informed philosophical theorizing is that cognition does not penetrate perception in this way.

For human vision, audition, and so on, there is a dominant set of norms, a convergence on colour discrimination, tone discrimination, and so on. And once developed to maturity, perceptual systems work rapidly and with no person-level effort, and most typically provide accurate representations. This has encouraged many theorists to think that perceptual systems must be biologically hard-wired, processing limited, modality-specific classes of information. Input from other non-sensory parts of the mind would undermine the observed cross-perceiver convergence on discrimination, speed, and objectivity and so, by inference to the best explanation, perceptual systems must be largely functionally independent. Therefore, perception is cognitively *im*penetrable. Or at the very least, our default should be scepticism about cognitive penetration *actually* occurring.[1]

The contrary claim has seen multiple definitions. One challenge is that given a particular case, extant, distinct definitions yield different verdicts: one definition will say of a case that it *is* cognitive penetration, and another definition, that it is not. This kind of theoretical cross-talk thwarts progress. One method to capture a unified meaning across extant definitions appeals to the alleged consequences of the possible phenomenon.

First, an interest in cognitive penetration traces back to a concern about possible theory-ladenness—whether background theoretical commitments affect perception in such a way that the rational role of perception in scientific theory choice is threatened [Hanson 1958, 1969; Kuhn 1962]. This consequence generalizes to a second one concerning everyday, empirical

[1] The two standard defenders of this kind of view are Jerry Fodor [1983, 1984, 1985, 1988] and Zenon Pylyshyn [1980, 1984, 1999] And, with varying degrees of commitment, there have been numerous more recent examples. See Burnston and Cohen [2015]; Gross et al. [2014]; Machery [2015]; Raftopoulos [2001].

knowledge. If one sees or otherwise perceives in ways infected by what one wants or antecedently believes, then the knowledge providing role of perception is threatened (or at least affected)[Lyons 2011; Siegel 2012, 2013; Stokes 2012; Silins 2016]. The final consequence concerns mental architecture, most centrally, the modularity theory of mind. Strong versions of this theory countenance perceptual systems as *informationally encapsulated*. Information from cognitive systems—reasoning, decision making, beliefs, goals—remains partitioned off from perceptual systems. Thus evidence for cognitive penetration—a violation of informational encapsulation—threatens revision of this model of mental architecture.

These consequences yield a debate-neutral working characterization of cognitive penetration. The phenomenon of interest is a cognitive-perceptual relation that implies one or more of these three consequences. Cognitive penetration involves some non-trivial cognitive effect on perceptual experience, where this effect plausibly results in consequences either for the epistemic role of perception (in the scientific or the everyday case), or the architecture of perception. Notice how the central emphasis here is on perceptual *experience*, by contrast to subpersonal perceptual *processing*.[2] Notice also how this allows for a variety of mental phenomena that "count as" phenomena of interest, since distinct phenomena may imply distinct consequences.[3]

The value of the working consequentialist characterization is illustrated by how it engages with standard sceptical replies to alleged cases of cognitive penetration. There are a number of standard replies—the judgment interpretation, memory interpretation, intra-perceptual

[2] In this respect, the primary emphasis of the current project differs from that of a few of the only other extant discussions of attention and cognitive penetration. The central emphasis here is on conscious perceptual experience. By contrast, Raftopoulos [2001], Cecchi [2014], and Wu [forthcoming] centrally emphasize whether cognition can affect computational processes of vision via attentional modulation.

[3] This consequentialist approach, defended in Stokes [2015], need not be assumed for the analysis given in this paper. As discussed in §2 below, there is an alternative way of analyzing attention-involving cases of perception such that they involve a direct, internal relation between cognition and perception and thus are, plausibly, cognitive penetration.

interpretation (see Macpherson [2012]; Stokes [2013]; Stokes and Bergeron [2015]). The interest here is in the *attention-shift interpretation*.

Jerry Fodor gives the attention-shift interpretation in numerous places [1983; 1984; 1985; 1988](see also Pylyshyn [1999: 358, 364]). Suppose one alleges that one's ability to see the duck-rabbit image sometimes as a duck and sometimes as a rabbit is a case of cognitive penetration (see Figure 1)[Churchland 1988]. The sceptic can grant that here cognition is affecting perception, but only through shifts in spatial attention—that is, changing the focal space to which one, in this case, visually attends. Fodor writes, "one squints to make things look sharper; one cups one's hand behind one's ear to make them sound louder, etc. It doesn't begin to follow that auditory and visual acuity are cognitively penetrable" [Fodor 1988: 190-1].

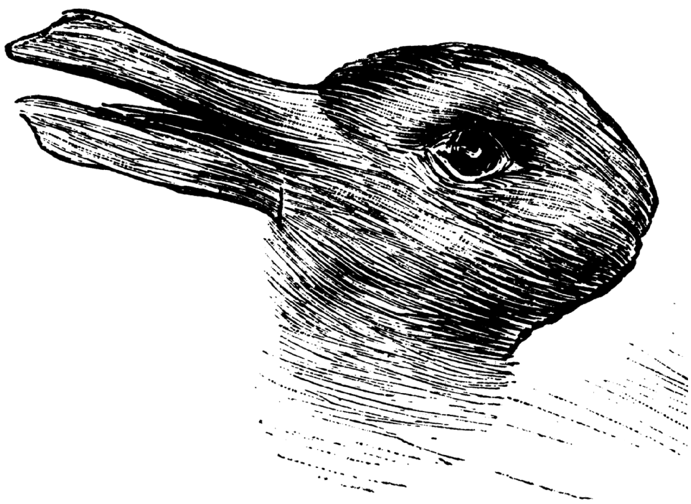

Figure 1

Appeal to consequences is instructive here. Fodor's reasoning might concern either of two possible features of attention shifting. First, shifts in attention are often *actions* in the philosopher's sense. Fodor describes the duck-rabbit case in precisely this way: cognitive states cause action, and those actions affect what information reaches sensory receptors, which in turn affects perception. This bears no important consequences for the epistemology or alleged modularity of perceptual systems or experience. Second, agency aside, shifts in attention involve a change in the visual spatial array of attention. Here cognition changes the visual spotlight or focus and accordingly changes the information available for perception. Again Fodor can charge that this bears no consequence for the epistemic role of perception, nor any obvious implication for the alleged modularity of perceptual systems. Instead, it is just a change in the input to those systems, by contrast to an encapsulation-violating effect on the computational processing of those systems.

Grant for the sake of argument that if the attention-shift interpretation best explains a case then the sceptic is right, the case in question is not cognitive penetration. But what is the scope of the attention-shift interpretation? Sceptics like Fodor and Pylyshyn assume that a shift in spatial attention is the only relevant way that attention might be involved in cognitive-perceptual phenomena. This assumption is mistaken.

2. How can attention mediate cognition and perception?

How may attention count against, versus be neutral with respect to (if not count for), explaining a phenomenon as cognitive penetration? The standard attention-shift interpretation countenances a scenario where a perceiver has certain cognitive states that cause a shift in attention and that shift clearly changes what the perceiver sees. This scenario is captured by a simple, mental causal schema:

(a) Cognitive state ➔ *Attention-shift* ➔*Perceptual experience*

One way to understand attention-shifting is as an intentional bodily action. Second, and non-exclusively, attention-shifting can be understood as a shift in spatial attention: a change in the spatial visual field attended, to the focal spotlight. And again, let's grant that the sceptic is correct to claim that if a case is best explained in terms of (a), then it is not a case of cognitive penetration.[4]

Are all attention-involving, cognitive-perceptual phenomena best described in terms of schema (a)? To motivate a negative answer to this question, affirmative answers for each of the following questions must be given:

> Q1: Can attention (or attentional mechanisms) operate in a non-agential way?
>
> Q2: Can attention change without changes in spatial attention?
>
> Q3: Can these non-spatial attentional mechanisms be influenced or partly driven by cognitive states?
>
> Q4: Do these non-spatial attentional mechanisms influence conscious perceptual experience?

In response to Q1, it should be obvious that attention can be guided without one's trying or intending to do any guiding. Some objects or events grab attention—a snake popping out of the brush, one's name mentioned on the opposite side of the room, a sharp pain in one's foot. And this comports well with the most dominant, general categories of scientific theorizing about attention. Bottleneck theories [Broadbent 1958], Feature-binding theories [Triesman and Gelade 1980], and Competition-based theories [Desimone and Duncan 1995] are broadly unified by the following thread: attentional mechanisms serve to *select* perceptible features, objects, and events, and often automatically. This suffices to answer 'yes' to Q1.

Q2 asks, can attention change without changes in spatial attention? Attention is used, intentionally and non-intentionally, as a way to search, recognize, and inspect one's environment. One clear way this occurs involves focused attention to some region of space in the visual field,

[4] It would be more accurate, then, to replace 'attention-shift' in schema (a) with either 'focal attention shift' or 'intentional/spatial attention shift' (and the schema, again, should be read that way). However, schema (a) has been termed in keeping with the standard terminology in the relevant literature, which widely discusses the 'attention shift interpretation'. Thank you to an anonymous referee for emphasizing this qualification.

for example, the region where one might expect to find the object of one's visual search. This is one standard account of spatial attention—the "spotlight of attention"—and it's the account assumed by the Fodorian sceptic. But it should be unsurprising that attentional selection can instead involve enhanced sensitivity to features of one's environment (like colour or shape), or to whole, bound objects in one's environment. And this sensitivity can be enhanced in *behaviourally relevant* ways. Over the last few decades, there has been substantial empirical research on both of these last two mechanisms of *selective attention*: on *object-based attention* (OBA) and *feature-based attention* (FBA).

The earliest work on object-based attention comes from Neisser and Becklen [1975]. Experimenters asked subjects to view a screen with two distinct, spatially overlapping videos. In one version of the study, one video depicted two sets of hands playing a hand-slapping game, and the second video depicted a group of people passing around a basketball. When subjects were tasked in a way specific to one video—for example, to count the number of basketball passes—they would miss events taking place in the other video. The videos in this experiment are spatially superimposed and so if attention was operating just on the basis of a spatial position (or "spotlight"), then subjects should not be blind to changes in those positions. But in fact, robust *inattentional blindness* occurs. This is best explained by appeal to an attentional mechanism that responds to objects and does so independently of spatial attention (see Scholl [2001]).

In visual search, attentional selection is sometimes guided by relevant features of the environment. Suppose you are searching a crowded baggage claim area for a missing piece of bright pink luggage. Feature-based attention operates in such a way that when scanning this visual array, pink-ish features of the array are favoured for perceptual representation at the neglect of other features. Pink-ish features enjoy a brief pop-out effect to aid in search. Importantly, like OBA, FBA can operate independently of spatial attention: FBA will highlight relevant features both "within and outside the current spotlight of attention" [Treue and Martinez-Trujillo 2007: 175](See also Scolari et al. [2014]; White and Carrasco [2011].

So, in answer to Q2: both OBA and FBA can function, selecting objects or features, without any change in spatial attention.[5] Note further that these selection mechanisms can operate in non-agential ways. Q3 then asks if these mechanisms can be modulated by cognition and, Q4, whether these mechanisms function in such a way to influence conscious perceptual experience.

Consider *Where's Waldo?* puzzles: overwhelmingly busy visual scenes where one is tasked to find Waldo, who is typified by matching red and white striped sweater and cap, and round spectacles. This cognitive-perceptual situation is plausibly interpreted as follows. The practiced *Where's Waldo?* puzzler has a number of relevant cognitive states: a belief that there is an object in the scene with Waldo's typical features, an expectation to find items with those features, an intention to find those typical features. FBA then functions in such a way that behaviourally relevant features (red and white striped items) are rapidly selected at the expense of non-relevant features. This suggests a 'yes' to Q3. This cognitively enhanced activity in FBA then results in or amounts to enhanced perceptual representation—pop-out—of some features at the neglect of others. This is an effect on conscious perceptual experience (a 'yes' to Q4).

In this case, cognitive states like belief and intention affect conscious experience, by modulating intermediate (but non-agential, spatially independent) selective attention. Contrast this with the naive *Where's Waldo?* viewer, tasked to simply inspect the same visual scene. This subject will not (immediately) form the relevant cognitive states, will not enjoy FBA-driven search and, finally,

[5] Another way to mark this distinction is in terms of focused vs. distributed attention. Attention can be focused on one region or space. Or, attention can be distributed across features or objects within a region of space. Putting the current points in these terms, distributed attention (to features or objects) can change without any change to the focus (the space) of attention.

will not enjoy the resulting, enhanced perceptual representation. Whatever one says about this anecdotal case, what's crucial to note is that the sceptic cannot invoke the attention-shift interpretation in response. A case of this kind (involving the non-naive viewer) does not involve shifting (intentionally or not) attention from one space to another. Instead, it is plausibly characterized by an alternative causal schema:

(b) *Cognitive state* ➔ *Non-agential selective attention*➔ *Perceptual experience*

Now consider schema (b) in light of three recent, related experimental studies.

Moores et al. [2003] found that object representations *and* associated cognitive representations influence attentional selection in visual search tasks. The basic experimental structure went as follows. Subjects were placed before a computer screen in a dimly lit room, viewing the screen at a distance of roughly 57cm. Subjects were cued with an instruction for visual search, for example the target item might be MOTORBIKE. Subjects would then see a fixation cross, followed by time variable presentation (sometimes as brief as 100ms) of an array of four images. Sometimes the target item was present and in other trials absent. Additionally, in some experimental trials, amongst the visual array was a semantically related but visually non-resembling image (a HELMET image when the target was MOTORBIKE). Subjects would then receive a masking screen, or a fixation cross, and then might be tasked to report whether the target image was present or absent, and/or be asked to make a forced choice concerning which of two items was recently seen. (For illustrations of trial structures, see Moores et al. [2003: 183, 185]).

Images semantically related to the search target were better recalled and identified, and these images sometimes hindered target identification and reduced recall of unrelated distractor images. These researchers hypothesize that selective attention is being captured in a way independent of the search task but that depends upon information about relations of meaning, stored in long term memory.

The same researchers found that semantically related images modulated saccadic eye movement.[6] Saccades were more typically directed at images semantically related with the explicit search target, at the near exclusion of unrelated distractors. And in trials where a related image was present in the visual display, the probability that the first saccade went toward the target is significantly reduced, and by contrast to control trials with no semantically related image where initial saccades are dominantly towards the search target image (the mean onset latency of the first saccade was 306ms)(see also Belke et al. [2008]; Gazzaley and Nobre [2012]). This is rapid, non-deliberate selection of attention that is sensitive to semantic relationships.

In a second set of studies [Meyer et al. 2007], subjects again performed a visual search over a time-limited visual array. The relevant difference is that on some trials, included in the visual array was an image of an object with an homophonous name. For example, if the target item was a BOW (weapon), on some trials an image of a BOW (ribbon) was present. The related images here are neither visually resembling, nor semantically related (the respective terms for a bow weapon and a bow of ribbon stand in no semantic relations). In line with the Moores et al. results, Meyer et al. found that when present, images with homophonous names (shared with the target) more quickly receive attention, are better recalled and identified, and tend to slow overall response time. Here again eye tracking studies corroborate: the first saccade after display onset (onset latency averaging between 211 and 220 msec) was more likely to go to the related

---

[6] Humans produce an average of 3-5 saccades per second, and saccades take an average of 100-200ms. Saccades can be voluntary, as when one intends to shift one's gaze to another part of the room. But they also occur involuntarily or "reflexively". These movements are ballistic: once a saccade is initiated (voluntarily or reflexively), it cannot be adjusted mid-saccade (e.g. if the target moves).

(homophonous) image than unrelated distractors, and the target was less likely to remain fixated if there was a related image present in the visual array.

Finally, these studies have been corroborated by neurological studies using electroencephalographic (EEG) recordings. In these studies, Telling et al. [2009] used measurements of N2pc amplitudes. The N2pc is an event-related potential (ERP), and is a standard measure for fairly rapid activity in response to visual stimuli, typically around or before 200ms. In brief, what these researchers found was that semantically related (but visually non-resembling) distractor items affected the magnitude and onset latency of this EEG component. This cortical activity is fast, and so it is improbable that it is attentional selection consistently done by, in any intentional way, the perceiver. Here it is worth stressing some relevant features of N2pc research. First, using MEG neuroimaging techniques, researchers identify the N2pc component as largely correlating with activity in extrastriate cortex, including visual area V4 [Hopf et al. 2000]. Activity in the latter visual cortical area has long been taken to be modulated by (or correlated with) selective attention [Moran and Desimone 1985]. Furthermore, recent research suggests that N2pc amplitudes are *not* correlated with shifts in spatial attention [Kiss et al. 2008]. Thus semantic information stored in long term memory appears to be modulating rapid neural processing that occurs, by our best neuroscientific models, *in* the visual system.

Now recall, Q3 asks: Can these non-spatial (selective) attentional mechanisms be influenced or partly driven by cognitive states? What all of these experimental studies converge on—from behavioural to eye-tracking to neurological studies—is that non-targeted but linguistically related images have a significant effect on visual search. The difference made by semantically and homophonously related images requires explanation. The methodology used, including the use of the N2pc EEG component, suggests that the difference is not well explained by shifts in spatial attention. Since representation of linguistic relations is not merely (or even) perceptual, even if “low-level” or “fast” or even “encapsulated”, the difference is not explained as a merely perceptual (that is, non-cognitive) one. And finally, the eye-tracking and EEG results suggest that the difference is not explained as a mere post-perceptual difference in judgment or memory. Instead the data suggests important effects both at the level of saccadic eye movement, and in extrastriate cortical activity, both of them part of visual processing if anything is.

As Meyer et al. describe it, the “data suggest that there is sufficiently rapid access to conceptual information from distractors for this information to influence the first fixations made during search” [2007: 715]. Recall that these first saccades averaged (in these studies and the Moore et al. [2003] studies) an onset latency between 200 and 306ms. It is not just the fact that the effect is rapid that is important, but the fact that linguistically related distractors are both task-irrelevant and visually non-resembling to the relevant target images. In this regard, saccadic movements to distractors are not deliberate but instead involve an “automatic" spread of attention [Meyer et al. 2007: 710][7]

This latter explanation comports well with an explanation in terms of selective attention, with emphasis on feature-based attention. Typically, researchers of FBA focus on low-level features like colour or motion, and how selection of these features may be modulated in behaviourally relevant ways. However, the data in these studies suggest that linguistic features (for example, that the image stands in a semantic relation with the search target) are operative, such that features (or objects) associated with those linguistic features are somehow enhanced. When searching for a MOTORBIKE image, features of semantically related images (like those of a HELMET image) are deemed relevant by the cognitive system and, accordingly, are selected for

[7] One might object that these studies are just yet another bit of evidence for semantic priming. Here are two additional qualifications against that rejoinder. First, the effects here are not all behaviourally enhancing, as priming effects often are (see Moores et al. [2003: 187-8]; Meyer et al. [2007: 714]). Second, the effects here are not, as priming is often defined, (mere) effects on memory. This is precisely the claim to be defended in reply to Q4.

attention at the cost of unrelated distractors. In brief, this would be working memory somehow modulating feature selection in favour not of relations of visual similarity to the target, but in favour of linguistic relations. This is an extension of the standard model of feature based attention, but it gains plausibility from the data for the simple reason that it best explains that data.

Q4 asks: do these non-spatial (selective) attentional mechanisms influence conscious perceptual experience? As forms of selective attention, FBA and OBA can affect perceptual salience without the perceiver performing any action to bring about the changes, and without any change to or manipulation of features of the environment. These perceptual changes can take a number of forms. Most simply, a feature or object that pops out is experienced sometimes at the exclusion of, or longer than, other stimuli in the array. Selective attention also effects changes in perceptual organization. Consider the Rubin Goblet (see Figure 2). Even holding the stimulus fixed in your spatial attentional spotlight, you experience a phenomenal difference between seeing the image as a goblet and seeing the image as two faces in a face-off. This is a perceptual change that you can deliberately bring about. But it is entirely plausible that selective attentional mechanisms, when operating non-agentially, bring about similar changes: changes in what features or objects are figure(s) versus ground, and what features or objects are focally versus peripherally represented.

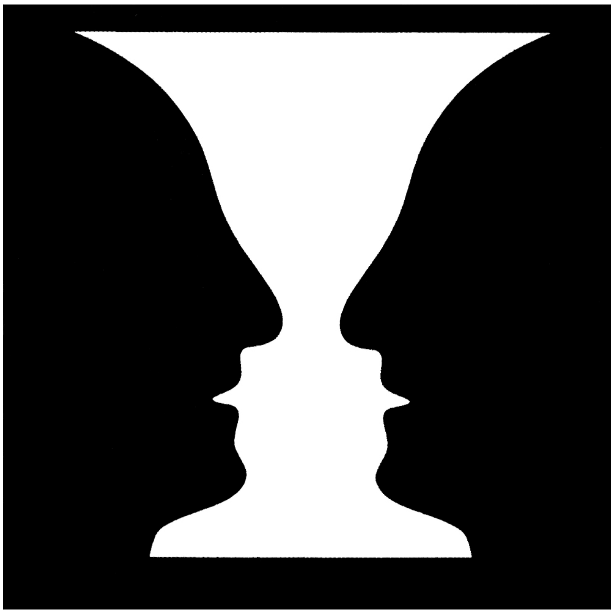

Figure 2

Recent theorists have suggested that attention serves to *structure* or organize conscious experience [Watzl 2011; Wu 2014, Ch. 4]. This observation traces back to gestalt psychology and continental phenomenology. It is clearly articulated by Sartre. "When I enter this cafe to search for Pierre, there is formed a synthetic organization of all the objects in the café, on the ground of which Pierre is given as about to appear. […] [I]f I should finally discover Pierre, my intuition would be filled by a solid element, I should be suddenly arrested by his face and the whole cafe would organize itself around him as a discrete presence" [Sartre 1943: 9-10]. Should Sartre find, attentionally select, Pierre's face, it is clear that Pierre's face would take up the central, focal position in Sartre's conscious experience. More interestingly, it is around this one object or feature—Pierre's face—that Sartre's conscious visual experience is then organized. So, it is crucial to note not just that there is a centre or a priority space of attention, but that there are various other features of the perceptual array that are organized by relation to that centre. These components are no less a part of the phenomenology of experience. And the selections that attention makes determine how all of these components are structured.

The explanation in reply to Q3 regarding the discussed experimental studies extends to a reply to Q4 as follows. The semantically and linguistically related items receive attention at the exclusion of unrelated distractors, or for periods longer than those distractors (as suggested by the eye-tracking studies). This is the simplest perceptual effect, showing up in the relevant subjects' performance on recall tasks. It is also plausible that these attentional selections influence the structure of conscious experience in the ways just described. For example, the targeted MOTORBIKE image might take the centre of attention, with the HELMET image "pulling" dominantly at the periphery of experience. And this experience enjoys a distinctive phenomenology: it feels different to have something peripheral pulling attention from your focus vs. a case where attention is perfectly focalized. These are attention-modulated perceptual differences that causally depend upon the cognitive, linguistic representations stored in long term memory.

Brief reminder of the dialectic is useful here. The central question is: how may attention count against, versus be neutral with respect to (if not count for), explaining a phenomenon as cognitive penetration? This question can be approached by asking whether all attention-involving cognitive-perceptual phenomena are best explained in terms of the attention-shift interpretation (causal schema (a))? To answer 'no' to this last question, an affirmative answer must be motivated for questions Q1-Q4. This has now been done, and by appeal to both anecdotal and experimental evidence. This analysis grounds the second causal schema:

(b) *Cognitive state* ➔ *Non-agential selective attention*➔ *Perceptual experience*

The very same evidence and considerations that motivated affirmative answers to Q1-Q4, suggest that we have compelling evidence for actual psychological phenomena that are best described in terms of schema (b). It should be clear by this stage of the analysis that a phenomenon that fits (b) would not, for inclusion of a central role of attention, count *against* that phenomenon being cognitive penetration. The final step in the analysis is to argue that any phenomenon that is best described in terms of (b) is best explained *as* cognitive penetration.

Cognitive penetration is often theorized as involving a *direct* cognitive effect on perception. Pylyshyn [1999] and Fodor [1988] at least implicitly assume this in their claim that cognitive-perceptual effects involving active shifts in spatial attention do not count as cognitive penetration. Cases involving spatial attention shifts fail to satisfy what recent theorists have called a *vehicle* criterion, which requires of genuine cognitive penetration that perceptual processing must draw information or representations directly from cognitive systems [Raftopoulos and Zeimbekis 2015: 27]. More weakly, some have allowed for *indirect* connections [Macpherson 2012], while also maintaining that the connection must be *internal* and *mental* [Stokes 2013]. Attention-shift cases best described in terms of causal schema (a) do not count as cognitive penetration by failure of all of these criteria. But on one plausible analysis, this is not true of phenomena that are best described in terms of causal schema (b).

First note that a phenomenon that fits (b) does not involve (any relevant) intermediate action, nor does it involve a shift in spatial attention. Recall further that neural correlates for selective attention are found in visual areas such as V4 of the extrastriate cortex. Moreover, an observable signature of selective attention mechanisms like feature-based attention and object-based attention is saccadic eye movement patterns. Many researchers would count these physical activities—V4 activity and involuntary saccadic eye movement patterns—as part of visual processing and, therefore, count (some of) selective attention as *part of*, rather than antecedent to, visual processing. Finally, there is principled reason to carve up the space in this same way. What these selection mechanisms do is change the salience of a perceiver's immediate environment, in effect changing what the eyes pick up. And this is part of what perception itself does. If this analysis is correct, then a phenomenon that is best-described in terms of (b) involves a direct cognitive (and certainly an internal, mental) effect on perception. Since the intermediary selection mechanisms as schematized in (b) are part of perceptual processing (and as argued above, have an influence on perceptual experience), cognitive modulation of those mechanisms *is* cognitive modulation of perception. So, a phenomenon that fits (b) is cognitive penetration.[8]

One rejoinder says that the phenomena discussed here *are* or involve instances of spatial attention. For example, behaviourally sensitive changes in feature selection could enhance sensitivity just to those spaces where the features are present, and perhaps one would want to call this "spatial attention". However, this would not amount to an instance of schema (a), since 'attention' in that schema has been defined in terms of attentional spotlight or focus, sometimes actively controlled by the agent, and this follows the analyses given by Fodor, Pylyshyn, and others. So, *if* one theorizes some of these mechanisms as nuanced forms of spatial attention, and those mechanisms are cognitively sensitive and influence perception, this would just be an additional example of attention-mediated cognitive effects on perception, not dismissed as an attention-shift.[9]

A second rejoinder insists that selective attention and perception are, always, separate mental processes. Such a theorist would then claim that a phenomenon fitting (b) is not cognitive penetration for failure to satisfy the directness (or vehicle) criterion described above. However, here it is important to ask why directness is supposed to be required for cognitive penetration.

---

[8] Similarly, Mole [2015] argues against the claim that attention serves only as a gate-keeper between cognition and perception, at most changing the input to separable perceptual processes. Mole adduces recent empirical work by Kravtiz and Behrmann [2011] that suggests that colour, shape, and object perception, as well as visually detected orthographic categories, are not downstream from attention, but instead are richly intertwined with covert attentional selection, the latter of which is apparently sensitive to learnt cognitive information. Mole buttresses this analysis (following his own philosophical theory of attention [2011]) with an "integrated competition" model of attention, which "has no room for the idea that attention stands as a partition between cognition and perception" [Mole 2015: 233].

[9] Thanks to an anonymous reviewer for criticism on this point.

One diagnosis is this: a direct cognitive effect on perception is the best evidence that we could have for a cognitive-perceptual phenomenon of important consequence. Direct cognitive influence on perceptual processing would provide the clearest threat for the modularity of sensory systems, and imply important consequences for the knowledge providing role of perceptual experience. Suppose that this is correct. It only implies that the directness criterion is an evidentialist criterion in a weak sense: direct effects are, in principle, the best or most easily identified evidence for cognitive penetration. This alone does not imply that direct effects are the only types of relevant effect, nor that evidence for directness is the only evidence of relevance. Indeed, if directness is at most an evidentialist (or operationalist) criterion, then to insist that directness is a necessary condition for the phenomenon is to set the bar for cognitive penetration in a way that is unprincipled, and in favour of the sceptic. Instead, it seems, it is not the directeness of the cognitive effect that is essential but the consequences.

The question then becomes: what does a phenomenon fitting (b) imply about the consequences? Recall that a modular theory of perceptual systems maintains that perceptual processing is informationally encapsulated. A phenomenon that fits (b) would plausibly violate this condition, since intermediary selective attention is sensitive to the perceiver's learned cognitive states. This selectivity, of features or objects, directly influences the informational processing of perceptual systems. And as hypothesized above, this kind of attentional selection may operate non-agentially and independently of spatial attention. Therefore, perception is not informationally encapsulated relative to cognitive information. It is instead sensitive to learned individual and cultural differences. This is especially perspicuous in the empirical studies discussed above, where perceptual processing—evidenced by behavioural, eye-tracking, and EEG results—varies with linguistic concepts or knowledge. These results suggest that informational processing is sensitive to something more than only "purely" visual features of the external stimuli. So, even if the causal process runs from cognition through attention to perception, the standard modular cognition/perception architecture is called into question.

Now consider the epistemological consequence (concerning theory-ladenness or general empirical knowledge). Fodor was clear that modular systems are epistemically preferable. Perceptual processing *independent* of background beliefs, goals, and intention will more rapidly and reliably deliver objective representations about the immediate environment [Fodor 1983, 1985, 1988]. And so as these cognitive influences go up, the relevant perceptual independence goes down. Recall that the answer to Q4 above was that selective attention can influence what (and for how long) objects or features are experienced, what "pops-out" in experience or is salient, and how experience is structured. And recall that selective attention can affect perceptual salience and structure without the perceiver acting to bring about the changes, nor changing or manipulating input from the perceptible environment. These are genuine perceptual effects. Insofar as these effects are sensitive to background cognitive information, rather than just features of the immediate environment, the independence of perceptual representation is called into question. Thus, some epistemic consequence for perception follows.[10]

One might object: "Surely these effects are too minor to sound sceptical alarm bells. In the empirical studies discussed, for example, there were no massive errors made and no illusions suffered. So why should we be worried about the epistemic role of perception on the basis of evidenced phenomena that fit (b)?" An important lesson can be gleaned from this rejoinder. While it is important to note that some perceptual error, or loss in efficiency, did occur in the studies on linguistic effects, it is more important to note that the relevant epistemological consequence need not be pernicious. This is typically, and sometimes explicitly, supposed. But

[10] Thanks to Chris Mole for helpful suggestions on this point.

there is no principled reason to assume that cognitive penetration, should it occur, is bound to produce epistemic problems. In other words, the epistemic consequence can be stated neutrally. What matters is that cognition affects perception *in some important way* with respect to the knowledge-providing role of perception. In Fodor's terms, but contrary to his conclusion, epistemic independence of perception may not be necessary or even especially crucial to objective perceptual representation. And in fact, this kind of effect could be epistemically boosting, rather than downgrading (see Lyons [2011]; Stokes [2013, 2014]; Siegel [2013]; Vance [2015]).

One clear way to bring out this possibility is to consider expertise. Subjects in the Moores et al. [2007] and related experiments had to have some minimal linguistic expertise. Absent that expertise, there is no relevant perceptual effect, and accordingly no relevant effect on the perceptual judgments that subjects make about the visual field, nor an effect on performance time or accuracy. Thus linguistic expertise influences, it seems, perception in a way that in turn influences reports and doxastic commitment. There is rich experimental literature on experts in more specialized domains, worthy of brief mention at least for illustration. For example, radiologists reliably identify abnormalities in rapidly presented (200ms) chest radiographs. Eye-tracking studies corroborate these results by revealing substantially different scan patterns between experts and controls: expert radiologists make fewer total saccadic eye-movements and fixate more quickly on abnormalities [Drew et al. 2013; Evans et al. 2013]. These and many other cases of empirically studied perceptual expertise are candidate examples of epistemically enhancing cognitive penetration. Consideration of these cases should at least encourage caution about the assumption that the epistemic consequence of cognitive penetration must be pernicious (see Bukach et al. [2006] and Scott [2011]).

This concludes the argument that attention-involving cognitive effects on perception need not take the form of causal schema (a). Instead, there is good reason to think that phenomena described in terms of causal schema (b) are possible, and substantial empirical evidence that actual phenomena do fit this very schema. Finally, and in answer to the central question of this paper, these are plausibly attention-mediated instances of cognitive penetration.[11]

3. Concluding lessons

This is progress for the cognitive penetration debate. The framework offered here moves beyond a simple assumption of extant theorists. That assumption is captured by the following conditional claim: If attention is part of the causal explanation of a relation between cognition and perception, then that relation is not cognitive penetration. This conditional is true only if one grants the further assumption that the following causal schema is the only possible one.

(a) Cognitive state ➔ *Attention-shift* ➔*Perceptual experience*

But there is an additional important causal schema:

(b) Cognitive state ➔ *Non-agential selective attention*➔ *Perceptual experience*

[11] Wu [forthcoming] gives an analysis in the same spirit and comes to the same *general* conclusion as this one, but the emphasis and the mode of analysis are substantively distinct. Wu, like Raftopoulos [2011] and Cecchi [2014], takes violation of informational encapsulation to be an important marker for cognitive penetration. Wu's central question is whether background cognitive states directly influence the computational processing in visual structures. Wu argues, by appeal to neurological studies on Macaque monkeys and computational models, that intention modulates visual attention; and because intention is cognitive and attention is an "aspect" of vision, this counts as cognitive penetration. So Wu's analysis centrally concerns penetration of perceptual processing and directness, by contrast to the emphasis here on perceptual experience and theoretical and epistemic consequences. But like this paper, Wu partly emphasizes selective attention and explicitly notes the importance of epistemological consequences of intention-involving cases of cognitive penetration.

So the assumed conditional is false. There are other ways that attention could be involved in cognitive effects on perception. This is the first lesson.

The second lesson is that, if this analysis has been successful, any *actual* psychological phenomenon that fits schema (b) is cognitive penetration. And this changes the dialectical space. The question now becomes, in the context of thinking about attention and cognitive penetration, are there *any* actual phenomena appropriately described to fit schema (b)? The sceptic must motivate a 'no' in answer to this question. And he must do this in the face of the mounting evidence from perceptual psychology discussed here. Put strongly, in the context of thinking about attention and cognitive penetration, the sceptic's position is thereby undermined. The more cautious conclusion is that, at the very least, the burden of proof shifts to the sceptics, to those who deny the cognitive penetration of perception.

Acknowledgements

Earlier versions of this paper were given at Kansas State University; The Tanner Humanities Center-University of Utah; the Workshop on Cognitive Penetrability and Predictive Coding-University of Bochum, Germany; the Bled Philosophical Conference on The Intersection of Epistemology and Philosophy of Mind- Bled Slovenia; the Perception and Cognition: Top-Down Influences in Perception Workshop-University of Glasgow; and the Conference on Impure Perception, Berlin School of Mind and Brain-Humboldt University, Berlin. Thanks to all participants at these venues. I owe special thanks to Peter Fazekas, Fiona Macpherson, Chris Mole, and Jona Vance for discussion of these issues. Finally, thanks to the editor and two helpful, anonymous referees for this journal.

REFERENCES

Belke, E., G.W. Humphreys, D.G. Watson, A.S. Meyer, and A.L. Telling 2008. Topdown effects of semantic knowledge in visual search are modulated by cognitive but not perceptual load, *Perception & Psychophysics* 70/8: 1444–1458.

Bukach, C.M., I. Gauthier, and M.J. Tarr 2006. Beyond faces and modularity: the power of an expertise framework, *TRENDS in Cognitive Science* 10/4: 159-66.

Burnston, D. and J. Cohen 2015. Perceptual integration, modularity, and cognitive penetration, in *The Cognitive Penetrability of Perception*, ed. J. Zeimbekis and A. Raftopoulos, Oxford: Oxford University Press: 123-43.

Broadbent, D. E. 1958. *Perception and Communication*, Oxford: Pergamon Press.

Cecchi, A. 2014. Cognitive Penetration, Perceptual Learning and Neural Plasticity, *Dialectica* 68/1: 63-95.

Churchland, P. 1988. Perceptual plasticity and theoretical neutrality: A reply to Jerry Fodor, *Philosophy of Science* 55/2: 167-87.

Desimone, R. and J. Duncan 1995. Neural mechanisms of selective visual attention, *Annual Review of Neuroscience*, 18: 193–222.

Drew, T., K.E. Evans, M.L.H. Vo, F.L. Jacobson, and J.M. Wolfe 2013. What can you see in a single glance and how might this guide visual search in medical images?, *Radiographics* 33/1: 263-274.

Evans, K.K., D. Georgian-Smith, R. Tambouret, R.L. Birdwell, and J.M. Wolfe 2013. The gist of the abnormal: above- chance medical decision making in the blink of an eye, *Psychonomic bulletin & review* 20/6:1170-1175

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

Fodor, J. 1984. Observation Reconsidered, *Philosophy of Science* 51/1: 23-43.

Fodor. J. 1985. Precis of *The Modularity of Mind*, *The Behavioural and Brain Sciences* 8/1: 1-5

Fodor. J. 1988. A reply to Churchland's 'Perceptual plasticity and theoretical neutrality', *Philosophy of Science* 55/2: 188-98.

Gazzaley, A. and A. Nobre 2012. Top-down modulation: bridging selective attention and working memory, *Trends in Cognitive Science* 16/2: 129-35.

Gross, S., T. Chaisilprungraung, E. Kaplan, J. Menendez, and J. Flombaum 2014. Problems for the purported cognitive penetration of perceptual color experience and Macpherson's proposed mechanism,*The Baltic International Yearbook of Cognition, Logic, and Communication* 9: 1-30.

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

Hanson, N.R. 1969. *Perception and Discovery: An Introduction to Scientific Inquiry.* Freeman and Cooper.

Hopf, J.M., S.J. Luck, M. Girelli, T. Hagner, G.R. Mangun, H. Scheich, and H.J. Heinze 2000. Neural sources of focused attention in visual search, *Cerebral Cortex* 10/12: 1233-41.

Kiss, M., B.A. Goolsby, J.E. Raymond, K.L. Shapiro, L. Silvert, A.C. Nobre, N. Fragopanagos, J.G. Taylor, and M. Eimer 2007. Efficient attentional selection predicts distractor devaluation: ERP evidence for a direct link between attention and emotion, *Journal of Cognitive Neuroscience* 19/8: 1316–1322.

Kravitz, D.J., and M. Behrmann 2011. Space-, object-, and feature-based attention interact to organize visual scenes, *Attention, Perception, and Psychophysics* 73/8: 2434-47.

Kuhn, T. 1962. *The Structure of Scientific Revolutions*, Chicago, IL: University of Chicago Press.

Lyons, J. 2011. Circularity, Reliability, and the Cognitive Penetrability of Perception, *Philosophical Issues* 21: 283-311.

Machery, E. 2015. Cognitive Penetrability: A No-Progress Report, in *The Cognitive Penetrability of Perception*. ed. J. Zeimbekis and A. Raftopoulos, Oxford: Oxford University Press: 59-74.

Macpherson, F. 2012. Penetration of Colour Experience: Rethinking the Issue in Light of an Indirect Mechanism, *Philosophy and Phenomenological Research* 84/1: 24-62.

Meyer, A. S., E. Belke, A.L. Telling, and G.W. Humphreys 2007. Early activation of object names in visual search, *Psychonomic Bulletin & Review* 14/4: 710-716.

Mole, C. 2011. *Attention is Cognitive Unison: An Essay in Philosophical Psychology*, New York: Oxford University Press.

Mole, C. 2015. Attention and Cognitive Penetration, in *The Cognitive Penetrability of Perception*, ed. J. Zeimbekis and A. Raftopoulos, Oxford: Oxford University Press: 218-37.

Moores, E., L. Laiti, and. L. Chelazzi 2003. Associative knowledge controls deployment of visual selective attention, *Nature Neuroscience* 6/2: 182-89.

Moran, J., and R. Desimone 1985. Selective Attention Gates Visual Processing in the Extrastriate Cortex, *Science* 229/4715: 782–4.

Neiser, U., and R. Becklen 1975. Selective looking: Attending to visually specified events, *Cognitive Psychology* 7/4: 480-94.

Pylyshyn, Z. 1980. Computation and Cognition: Isssues in the Foundations of Cognitive Science, *Behavioral and Brain Sciences* 3: 111-32.

Pylyshyn, Z. 1984. *Computation and Cognition*, Cambridge, MA: MIT Press.

Pylyshyn, Z. 1999. Is vision continuous with cognition? The case for cognitive impenetrability of visual perception, *Behavioral and Brain Sciences* 22/3: 341-365.

Raftopoulos, A. 2001. Reentrant pathways and the theory-ladenness of perception, *Philosophy of Science* 68/S3: 187-99.

Raftopoulos, A. and J. Zeimbekis 2015. The Cognitive Penetrability of Perception: An Overview, in *The Cognitive Penetrability of Perception*, ed. J. Zeimbekis and A. Raftopoulos, Oxford: Oxford University Press: 1-56.

Sartre, J.P. 1943/1969. *Being and Nothingness*, trans. H.E. Barnes. London: Routledge.

Scolari, M., E.F. Ester, and J.T. Serences 2014. Feature- and Object-based attentional modulation in the human visual system, *The Oxford Handbook of Attention*, ed. A.C. Nobre and S. Kastner, Oxford: Oxford University Press: 573-600.

Scholl, B. 2001. Objects and attention: the state of the art, *Cognition* 80/1-2: 1-46.

Scott, L. 2011. Face Perception and Perceptual Expertise in Adult and Developmental Populations, in *Oxford Handbook of Face Perception*, ed. G. Rhodes, A. Calder, M. Johnson, and J.V. Haxby, Oxford: Oxford University Press: 195-214.

Siegel, S. 2012. Cognitive Penetrability and Perceptual Justification, *Nous* 46/2: 201-22.

Siegel, S. 2013. The Epistemic Impact of the Etiology of Belief, *Philosophical Studies* 162/3: 697-722.

Silins, N. 2016. Cognitive penetration and the epistemology of perception, *Philosophy Compass* 11/1: 24-42..

Stokes, D. 2012. Perceiving and Desiring: A new look at the cognitive penetrability of experience, *Philosophical Studies* 158/3: 479-92.

Stokes, D. 2013. The cognitive penetrability of perception, *Philosophy Compass* 8/7: 646-663.

Stokes, D. 2014. Cognitive penetration and the perception of art, *Dialectica* 68/1: 1-34.

Stokes, D. 2015. Towards a Consequentialist Understanding of Cognitive Penetration, in *The Cognitive Penetrability of Perception*, ed. J. Zeimbekis and A. Raftopoulos, Oxford: Oxford University Press: 75-100.

Stokes, D. and V. Bergeron 2015. Modular Architectures and Informational Encapsulation: A Dilemma, *European Journal for Philosophy of Science* 5/3: 315-38.

Telling, A. L., S. Kumar, A.S. Meyer, and G.W. Humphreys 2009. Electrophysiological evidence of semantic interference in visual search, *Journal of Cognitive Neuroscience* 22/10: 2212–2225.

Treisman, A. and G. Gelade 1980. A Feature-Integration Theory of Attention, *Cognitive Psychology* 12: 97–136.

Treue, S. and J.C. Martinez-Trujillo 2007. Attending to Features inside and outside the Spotlight of Attention, *Neuron* 55: 174-6.

Vance, J. 2015. Cognitive Penetration and the Tribunal of Experience, *Review of Philosophy and Psychology* 6/4: 641-63..

Watzl, S. 2011. Attention as Structuring of the Stream of Consciousness, in *Attention: Philosophical and Psychological Essays*, Eds. C. Mole, D. Smithies and W. Wu, New York: Oxford University Press: 145–73.

White, A. and M. Carrasco 2011. Feature-based attention involuntarily and simultaneously improves visual performance across locations, *Journal of Vision* 11/6: 1–10

Wu, W. 2014. *Attention*, London: Routledge.

Wu, W. forthcoming. Shaking up the mind's ground floor: The cognitive penetration of visual attention. *Journal of Philosophy*.