ORIGINAL ARTICLE

[FINAL PREPRINT: PLEASE CITE PUBLISHED VERSION]

Recognition and the perception–cognition divide

[Greyson Abid, Department of Philosophy, University of California, Berkeley]

Correspondence

[Greyson Abid, University of California, Berkeley, Moses Hall, #2390, Berkeley, California 94720

Email: greysona@berkeley.edu]

Funding information

[Mabelle McLeod Lewis Memorial Fund]

[Recent discussions have fixated on the distinction between perception and cognition. How should recognition be understood in light of this distinction? The relevant sense of recognition involves a sensitivity to particulars from one's past. Recognizing the face of a familiar friend is one instance of this phenomenon, as is recognizing an object or place that one has viewed before. In this article, I argue that recognition is an interface capacity that straddles the border between perception and cognition.]

KEYWORDS

[recognition, familiarity, perception-cognition divide, dual aspect perception, informational encapsulation, long-term memory]

1 INTRODUCTION

Recent discussions in philosophy and psychology have focused on the distinction between perception and cognition.[1] This interest is not entirely new. Philosophers dating back to Aristotle have found the categories of perception and cognition to be theoretically fruitful ways of carving up the mind. Intuitively, the distinction is not difficult to appreciate. There is clearly *some* difference between seeing, touching, or tasting apple juice, on one hand, and thinking, reasoning, or making judgments about it, on the other. Intuition is partially vindicated by the success of scientific psychology, which readily employs such a distinction.

But how recognition should be understood in light of the distinction between perception and cognition. As a first pass (to be clarified shortly), the sense of recognition in question involves a sensitivity to particulars from one's past. Recognizing a familiar person (e.g., a colleague from work) is one instance of this, as is recognizing a place or thing that one has viewed before (e.g., a lake one visited as a child or one's jacket on a restaurant coatrack). It is not immediately apparent where recognition falls along the perception–cognition divide.

With a few notable exceptions, the topic of recognition (in the aforementioned sense) has been largely ignored in the philosophy of mind. This is surprising, given philosophers' interests in closely related matters, such as perceptual learning, imagination, and attention. One of the aims of this paper is to reignite philosophical interest in the topic. A natural starting point for any philosophical treatment of recognition involves determining what kind of mental phenomenon it is. Compare: if we discovered a new or long-forgotten chemical substance, a natural first step would be to determine its basic chemical kind.

A better understanding of recognition in light of the perception–cognition divide also helps clarify existing philosophical work on the significance of recognition. Strawson and Evans both emphasize the role of recognition in the identification of particulars. Strawson (1959) suggests that we would be unable to identify particulars at all if we could not recognize—or "reidentify"—them. Similarly, Evans (1982) argues that the capacity for recognition offers a way of identifying particulars that is more rudimentary than descriptive forms of identification. On Evans's view, recognition provides a means of identifying, for example, the Speaker of the

[1] See, for example, Block (2014), Firestone and Scholl (2016), Phillips (2017), Beck (2018), Mandelbaum (2018), and Green (forthcoming).

House of Representatives *without* entertaining a corresponding description in thought. Apart from its role in identification, Peacocke (1992) highlights that the capacity for recognition is a prerequisite for the possession of recognitional concepts, such as the concept "Lincoln Square". In considering the foundational role of recognition in identification and concept possession, it is natural to ask: What is needed to possess the capacity for recognition in the first place? Are only perceptual resources needed? Must cognitive resources be brought to bear? If so, which? By answering these questions, we more effectively spell out the implications of these works—an issue I return to in Section 7.

I begin by providing some clarificatory remarks concerning the perception-cognition distinction (Section 2) and recognition (Section 3). In Section 4, I offer a diagnostic tool for determining whether a mental phenomenon is at least partly perceptual. In Section 5, I use this diagnostic tool to make a case that recognition is at least partly perceptual. In Section 6, I draw on considerations relating to stimulus-independence and long-term memory to provide a qualified argument that recognition is also partly cognitive. The upshot is that recognition straddles the border between perception and cognition. I conclude in Section 7.

2 THE PERCEPTION-COGNITION DISTINCTION

A non-cursory discussion of the perception–cognition distinction would go beyond the scope of this paper. Nevertheless, it is worth making a few preliminary remarks. First, the distinction is neither exhaustive nor exclusive. It is not exhaustive in that certain affective states, such as undirected forms of depression or boredom, may be neither perceptual nor cognitive. It is not exclusive in that there may be capacities that share features of both perception and cognition. In the case of "borderline" capacities, it is indeterminate whether the capacity in question is perceptual or cognitive, analogous to the way in which it is indeterminate whether a virus is living or non-living (Block, 2014). In the case of "interface" capacities, the capacity in question is a (determinate) hybrid that is part perceptual and part cognitive, analogous to the way in which

a liger is a hybrid animal that is part lion and part tiger. Borderline and interface capacities are both compatible with a perception–cognition distinction. Indeed, I ultimately argue that recognition is an instance of the latter.

A hotly debated issue is whether top-down influences from cognition to perception impugn the hypothesis that perceptual systems (or "modules") are informationally encapsulated from central cognition. (A system X is "informationally encapsulated" from a system Y if and only if X cannot perform computations over information stored in Y. Less formally, such a system X cannot use system Y as an informational resource.) Clark (2013) and Lupyan (2015) claim that such top-down influences threaten the very existence of a perception–cognition distinction. If these theorists are correct, an attempt to understand recognition in light of the perception–cognition divide is fundamentally misguided.

Is this threat serious? There are two possibilities to consider. The first is that the hypothesis that perceptual systems are informationally encapsulated from central cognition withstands empirical scrutiny, rendering the threat empty. Perhaps all existing evidence against the hypothesis is methodologically flawed (Firestone & Scholl, 2016). Perhaps top-down influences, such as those mediated by attention, can be understood in a way that does not threaten the hypothesis (Quilty-Dunn, forthcoming). The second possibility is that the hypothesis is falsified. This possibility threatens the perception–cognition distinction *only if* one grounds the distinction in a form of informational encapsulation; so, if the hypothesis turns out to be false, one option—apart from abandoning the distinction—would be to find an alternative ground for it, such as a difference in representational format.

In any case, while I do presuppose that there is a perception–cognition distinction, my aim is not to characterize what grounds it. One benefit of this modesty is that it allows us to remain open to a range of considerations in assessing where recognition falls along the perception–cognition divide.

3 RECOGNITION

The lay concept of "recognition" is too ambiguous for philosophical discussion. In this section, I offer a more regimented characterization. After isolating the relevant sense of recognition (Section 3.1) and discussing its representational content (Section 3.2), I clarify how recognition (in the relevant sense) is distinguished from neighboring phenomena (Sections 3.3 to 3.4) and its relation to recollection and familiarity (Section 3.5), the personal-level (Section 3.6), and recognitional judgment (Section 3.7).

3.1 What is the relevant sense of recognition?

As noted earlier, my focus is on a sense of recognition that bears an intimate connection to one's past. I call this token-recognition, which is to be distinguished from what I call type-recognition. "Type-recognition" occurs when a subject groups a particular into a certain category on the basis of its appearance.[2] Without further qualification, this gloss is overly inclusive. A creature capable of color perception may group a sofa (that happens to be red) into the category of red things on the basis of its appearance, but few vision scientists would consider this is an instance of type-recognition. Similar considerations apply to certain other forms of low-level perception, such as depth or lightness perception. Why do we countenance face and object recognition, but not depth, color, or lightness recognition? A first pass answer is that in type-recognition a particular is grouped using a "sortal" category, roughly, a category expressed using a count noun that specifies a thing's (nominal or real) essence. This qualification is still not restrictive enough since type-recognition does not involve just any sortal category. Typically, the sortal categories are "basic-level" in Rosch's (1978) sense (e.g., DOG, CAR), striking a balance between more specific subordinate categories (e.g., GREYHOUND, ACURA) and more general superordinate categories (e.g., ANIMAL, VEHICLE).[3] A substantive question, which I do not discuss, is whether type-recognition is perceptual (see Mandelbaum, 2018).

Important for our purposes is that the type-recognition of some particular does not require any previous exposure to that particular, although it may require previous exposure to particulars

[2] Examples of particulars include people (better: individuals), places, and things.
[3] I use small capital letters to denote categories.

belonging to the relevant category. For example, I have been exposed to plenty of dogs in the past, so I have no trouble type-recognizing my cousin's newly adopted dog, even as it greets me for the very first time.

In contrast, "token-recognition" requires that an observed particular be registered as having been observed before. Here, "observation" should be construed broadly to include both direct and indirect observation. There are two reasons for this. First, an observed particular can be token-recognized even if it has not been *directly* observed before. For example, I might token-recognize Hugh Grant despite having never encountered him before in the flesh. Second, a previously observed particular can be token-recognized even if it is not *directly* re-observed. I might, say, token-recognize a friend in a painting. Nevertheless, in such cases there is *indirect* prior observation or re-observation through a representational medium. In the case of Hugh, the representational medium that allows for indirect prior observation is one of his films. In the case of my friend, the representational medium that allows for indirect re-observation is the painting that depicts them.

3.2 What is represented in token-recognition?

A token-recognized particular is *registered* as having been observed before. Talk of "registration" allows us to remain neutral regarding the difficult issue of what exactly is *represented* when one token-recognizes a particular. By way of example, consider the proposal that token-recognition involves representing an identity relation. On one version of this proposal, the content of such a representation might be expressed as "$that_1 F = that_2 G$". Here, "$that_1$" denotes a singular element that functions to refer to a *currently* observed particular, "$that_2$" denotes a singular element that functions to refer to a *previously* observed particular (similar to a memory based demonstrative), and "F" and "G" denote general elements that function to attribute (possibly identical) general features to the particulars referred to by the respective singular elements.[4]

[4] This approach extends Burge's (2010) framework.

This proposal raises thorny questions. Why think identity is the represented relation—as opposed to a relation of resemblance? And, supposing identity is represented, must the representing subject possess the concept of identity—as opposed to some nonconceptual analog of identity? There is a complex interplay between the answers to these questions and the topic at hand, namely, where token-recognition falls along the perception-cognition divide. For instance, if perception is nonconceptual in the sense that a subject can perceive that *p* without possessing the concepts that characterize *p* (cf., Byrne, 2005) and token-recognition falls on the perception side of the perception–cognition divide, the representation of identity involved in token-recognition must not require the representing subject to possess the concept of identity.[5] Even if we remain agnostic about issues concerning representation, it is worth remaining sensitive to these complexities.

3.3 Are type-recognition and token-recognition dissociable?

Clearly, type-recognition can occur without token-recognition. What about token-recognition without type-recognition? Certain examples are suggestive. I see a dog running through the trees at dusk without realizing it is a dog. I only discern that it is a medium-sized creature with brownish fur. Nevertheless, I am sure that it is the very same creature that was running through the trees a day ago. This is plausibly an instance of token-recognition without type-recognition. Of course, rejecting that token-recognition requires type-recognition is compatible with accepting that token-recognition is facilitated by type-recognition. I might have had an easier time token-recognizing the creature if I first grouped it into the category DOG. With these subtleties in mind, I tentatively accept that type-recognition and token-recognition are doubly dissociable and use "recognition" to refer exclusively to token-recognition in what follows.

[5] One might try to bypass these issues by sketching an alternative proposal. Perhaps a token-recognized particular is represented in a self-referential fashion using just one singular element that functions to refer to a currently observed particular. Building on the above proposal, the content of this representation might be symbolized as "that F which I have observed before". On this alternative, the previously observed particular is represented *as such*. Nevertheless, this proposal raises its own thorny questions of whether the representing subject must possess indexical concepts or the concept of observation. There is again a complex interplay between the answers to these questions and the issue of where token-recognition falls along the perception–cognition divide.

3.4 How is recognition different from short-term reidentification?

Recognition is distinct from various forms of short-term reidentification. Consider watching a cyclist weave through the streets. The cyclist may not always be in your direct line of sight as you track them. They may momentarily disappear behind a truck. Still, you continually track the cyclist through these brief occlusions. This phenomenon has received a great deal of attention among vision scientists. It is now well-established that visually tracking a particular through occlusion is subject to demanding spatiotemporal constraints. For example, visually tracking an object through occlusion is significantly impaired if the distance between the position of disappearance and reappearance of the object under occlusion is too large or if a small temporal lag is introduced during the occlusion period (Flombaum & Scholl, 2006). While our ability to recognize a particular certainly degrades over time, recognition is not subject to such highly circumscribed spatiotemporal constraints. Recognition is "spatiotemporally robust" in that a particular can be recognized even if it was last observed many months or even years before, perchance in a distant location. Just consider the possibility of recognizing a long-lost childhood friend while travelling in foreign country.[6]

Similar considerations apply to other forms of short-term reidentification. If I am shown an image of a face and am told that I will have to reidentify it a few seconds later, I may try to actively retain the image in visual working memory to complete the task. This strategy is only effective over small timescales since information held in visual working memory degrades within ~10 seconds (Zhang & Luck, 2009; Ricker & Cowan, 2010). Recognition is not subject to this constraint.

The claim that recognition is spatiotemporally robust is to be understood as a psychological generalization. Psychological generalizations (in general) are not exceptionless and usually contain hidden *ceteris paribus* clauses. It is for this reason that an intoxicated individual who can only hold one item in working memory is no counterexample to the psychological generalization that working memory has a capacity of ~4 items. The same holds

[6] Larzabal et al. (2018) provide an illustration of the spatiotemporal robustness of recognition by showing that individuals recognize pictures last observed 8-14 years earlier at a rate above chance.

true for many purported counterexamples to the generalization that recognition is spatiotemporally robust.

3.5 How does recognition relate to familiarity and recollection?

Registering that an observed particular has been observed before—in a manner not subject to the sorts of highly circumscribed spatiotemporal constraints mentioned above—*suffices* for recognition. More specifically, it suffices for a type of recognition known as "familiarity". Familiarity *does not* require the retrieval of contextual information from past episodes involving that particular, such as when or where one saw it. Walking down the street, I might cross paths with a familiar person I have met years before without being able to retrieve specific information about where or when I met them. James has an example of familiarity with a *postponed* retrieval of contextual information (or "recollection") in mind in the following passage:

> …I enter a friend's room and see on the wall a painting. At first I have the strange, wondering consciousness, "surely I have seen that before", but when or how does not become clear. There only clings to the picture a sort of penumbra of familiarity—when suddenly I exclaim: "I have it, it is a copy of part of one of the Fra Angelicos in the Florentine Academy—I recollect it there!" (James, 1890; cited in Yonelinas, 2010).

According to now widely accepted "dual-process" theories (Mandler, 2008; Yonelinas, 2010), separate processes underlie familiarity and recollection. The former involves a gradable signal detection process based in the perirhinal cortex. The latter involves an all-or-nothing memory retrieval process based in the hippocampus. Dual-process theorists hold that reidentification tasks *often* involve both underlying processes.

3.6 Is recognition personal or subpersonal?

On the views of Strawson, Evans, and Peacocke, it is a *subject's* capacity for recognition that is crucial for identification and concept possession. More generally, we attribute the capacity for recognition to a subject as a whole, not to neural or information-processing systems within that subject.[7] Of course, there may be subpersonal-level *correlates* of recognition. For instance, exposure to familiar stimuli result in reduced neural responses. This is known as "repetition suppression". Repetition suppression occurs in early visual areas, such as V2 (Huang et al., 2018). Nevertheless, V2 does not engage in "recognition" in any non-figurative sense. In Section 6, I discuss a model of recognition that employs subpersonal-level representations and operations on those representations.

Given this personal-level characterization, how are we to understand devices such as facial recognition systems? We might claim that these devices are only capable of recognition in a figurative sense—much like we claim that a thermostat only figuratively "knows" the temperature. Or, we might claim that the personal-level extends to such devices—much like we often extend the personal-level to non-human animals not generally regarded as persons. This, in turn, would allow us to hold that these devices are literally capable of recognition. Our intuition sits somewhere between these two extremes, varying depending on the configuration and behavior of the device in question along with contextual factors that raise or lower our standards for mentalistic attributions.

3.7 What is the difference between recognition and judgment?

It is crucial to distinguish recognition from recognitional judgment. A "judgment" is a doxastic attitude, similar to belief, that a subject takes towards some proposition *p*. Unlike non-doxastic attitudes (e.g., desiring), doxastic attitudes are governed by norms describing the conditions for epistemic justification and permissibility—what are often called "epistemic" norms. An example: If one is inclined to judge that *p* but discovers that the evidence for *p* is equivocal, all else equal, one ought to suspend judgment about *p*. This judgment would be epistemically impermissible.

[7] For a discussion of the personal-level versus subpersonal-level distinction, see Drayson (2014).

A "recognitional" judgment is a special instance of a judgment—one made when an observed particular is judged as having been observed before. In contrast, recognition does not require one to form any such judgement.

There are a few ways that recognition and recognitional judgment come apart. I might recognize a person on the street and only *later*, in an act of self-reflection, judge that I have seen them before. Here, my recognition of the person acts as a (partial) causal antecedent for my recognitional judgement; it also plays a normative role in justifying the recognitional judgment. Alternatively, I may see a stranger on the street and mistakenly judge that I have seen them before. In this case, successful recognition does not serve as a causal or justificatory precursor to my recognitional judgment. Finally, it is plausible that many instances of recognition occur without any recognitional judgment at all. Russell (1921) offers the example of a horse returning to its stable as one such instance. More generally, it is plausible that some animals and infants have the capacity for recognition but not recognitional judgment.

3.8 Recap

To summarize, the relevant sense of recognition is token-recognition, which occurs when one registers—in a spatiotemporally robust manner—that an observed particular has been observed before. Recognition is a personal-level capacity that does not require the retrieval of contextual information from past episodes involving that particular. Finally, recognition is distinct from recognitional judgment.

4 THE PERSPECTIVAL TEST-

Upon encountering a particular, we often exhibit a response that varies across certain changes in our perspectival relations to that particular. I call this type of variable response pattern *perspectival sensitivity.* With a few clarifications (Section 4.1 and Section 4.3), I argue that a

mental phenomenon is at least partly perceptual if it exhibits perspectival sensitivity.[8] My argument is inductive: Paradigmatically perceptual phenomena exhibit perspectival sensitivity (Section 4.2), whereas cognitive phenomena do not, at least not independently of perception (Section 4.4).

4.1 What exactly is a perspectival relation?

A non-exhaustive list of perspectival relations between some subject S and some particular *o* includes: The spatial position of *o* with respect to S along with *o*'s surrounding context and various conditions of the environment (e.g., illumination and background noise conditions). An exhaustive list would include all and only those relational features that determine S's *perceptual* perspective on *o*—what we often speak of as the viewing conditions (of *o* with respect to S). S's perceptual perspective on *o* is to be distinguished from S's *cognitive* perspective on *o*—the sort of perspective S has while entertaining a thought about *o* under some specific mode of presentation. Putting things this way makes it apparent that there is an element of circularity in appealing to perspectival relations—and, by extension, perspectival sensitivity—to determine whether some mental phenomenon is perceptual: Understanding these notions requires some antecedent grip on the perception–cognition distinction. For our purposes, this circularity is relatively unproblematic. Problems *would* arise if we were to use these notions to articulate the distinction between a perceptual and a cognitive perspective or the grounds for the perception–cognition distinction.

4.2 Perspectival sensitivity in perception

Let us consider some paradigmatically perceptual examples of perspectival sensitivity in which one's *subjective* response varies across certain changes in one's perspectival relations to a particular.[9] Consider first the example of viewing two same-sized trees from different distances (e.g., 50 and 100ft.). It is often said that this example illustrates the dual aspect of perception.

[8] Throughout, I have a broad extension of "phenomenon" in mind that includes states, events, processes, and capacities.

[9] In Section 5, I focus on examples of perspectival sensitivity involving *behavioral* responses.

One aspect corresponds to the "apparent" sense in which the nearby tree appears larger than the more distant tree. The other corresponds to the "objective" sense in which the two trees appear to be same size.[10]

The trees appear to be the same objective size because of the *constancy* of visual perception. *Ceteris paribus,* each tree appears stable in objective size across certain changes in the perspectival relations between the viewer and the tree, notably, certain changes in distance. (For brevity, I omit the *ceteris paribus* clause in what follows.) This is aptly called "size constancy". At the same time, the trees differ in apparent size because of the *perspectival sensitivity* of visual perception. The apparent size of each tree varies as one changes one's perspectival relations to the tree, specifically, one's distance from the tree. Thus, the tree viewed from 50ft. has a larger apparent size than the same-sized tree viewed from 100ft., occupying a greater portion of one's visual field.

Another example that demonstrates the constancy and perspectival sensitivity of visual perception involves color appearance.[11] Seeing a monochrome pink wall partially covered in shadows reveals both a uniform and variable color appearance. The uniform color appearance of the wall is made possible by the constancy of visual perception. This appearance stays constant across certain changes in one's perspectival relations to different portions of the wall, changes which alter viewing position, color context, illumination, and so on. This is called "color constancy". In contrast, the variable color appearance of the wall is made possible by the perspectival sensitivity of visual perception. This appearance does not stay constant across the aforementioned changes in perspectival relations to different portions of the wall; analogous to apparent size, variable color appearance is sensitive to one's perspectival relations to different portions of the wall.

The constancy and perspectival sensitivity of visual perception extends to other sense modalities as well. As one moves closer to a concert venue's speaker system, there is a clear

[10] This example is given by Peacocke (1983), Tye (2000), Noë (2004), Schellenberg (2008), and Lande (2018). Schwitzgebel (2006) is skeptical as to whether perception possesses a dual aspect. Although disputes concerning phenomenology are always difficult to resolve, there is ample evidence that lay perceivers can readily distinguish "apparent" from "objective" appearances, suggesting that the dual aspect of perception is not merely a philosophical construct (Wagner, 2006; Green & Schellenberg, 2018; see also Morales et al., 2020).

[11] This example is given by Noë (2004) and Schellenberg (2008).

sense in which the music seems to get louder, as evidenced by the fact that, for example, I must shout to be heard when near the speaker. This is compatible with it seeming to me that the actual *source volume* remains roughly the same. This happens because we can distinguish alterations in the perceived loudness of a sound of constant intensity. The perceived loudness of the sound is made possible by the perspectival sensitivity of audition. Perceived loudness varies across certain changes in distance to the source (e.g., the speaker system). The perceived intensity of the sound is made possible by the constancy of auditory perception. The perceived intensity of the sound does not vary across changes in distance to the source, at least within a certain range. This is somewhat confusingly called "loudness constancy" (Zahorik & Wightman, 2001).

4.3 The perspectival test

I contend that we can employ perspectival sensitivity as a diagnostic tool. According to what I call the "perspectival test", *a mental phenomenon is at least partly perceptual if it exhibits perspectival sensitivity.*

Three clarifications are in order. First, the perspectival test indicates a mental phenomenon is at least partly perceptual *in the absence of defeaters.* For example, olfactory states exhibit perspectival sensitivity. Burge, however, does not regard olfactory states as perceptual on the grounds that they are not *about* (in the intentional sense) distal features of the environment.[12] While I am skeptical of Burge's analysis here, I note that it does not conflict with the perspectival test since a failure to exhibit the requisite form of intentionality required for perception would qualify as a defeater.

Second, the perspectival test is modest. It only indicates that a mental phenomenon is *at least partly* perceptual. It does not specify whether the phenomenon in question is purely or only partly perceptual. What does it mean to be *partly* perceptual? It cannot simply amount to being causally influenced by perception; this condition is far too weak to be of any theoretical interest.

[12] Burge regards olfaction as non-perceptual because it does not result in "objectification." According to Burge, "*Objectification* is [the] formation of a state with a representational content that is *as of* a subject matter beyond idiosyncratic, proximal, or subjective features of the individual" (2010, p. 397). For discussion suggesting that olfaction does result in objectification, see Begby (2011).

Rather, a mental phenomenon is "partly perceptual" just in case it in part constitutively depends on some perceptual element (e.g., a perceptual mechanism or representation). "Partly cognitive" is defined similarly. In Section 6, I argue that recognition is partly cognitive in virtue of its partial constitutive dependence on long-term memory representations. This claim does not conflict with the perspectival test. Rather, it implies that recognition is partly perceptual and partly cognitive.

Third, the perspectival test does not require that all perceptual phenomena exhibit perspectival sensitivity. Hallucinations are arguably perceptual, at least according to non-disjunctivists who claim that hallucinations and non-hallucinatory perceptual experiences belong to the same fundamental kind. Yet, hallucinations cannot always exhibit perspectival sensitivity since we are not perceptually related to any particulars in instances of "pure" hallucination. This is not in tension with the test. Passing the test is sufficient but not necessary for counting as at least partly perceptual.

Similar considerations apply to certain forms of imagination. Suppose I see a tree from 50ft away, turn so I no longer see it, start to imagine it, and move 100ft away from it—in that order. As I imagine the tree, its apparent size does not vary across *actual* changes in my distance from it, though it may vary across *imagined* changes in distance. While this form of imagination fails to exhibit perspectival sensitivity, this does not preclude the possibility that it is at least partly perceptual—a conclusion we may reach in light of certain neural, phenomenological, or representational similarities between imagination and perception.

4.4 Assessing the adequacy of the perspectival test

Do cognitive phenomena ever exhibit perspectival sensitivity? Of course, there are cases in which we form *perceptually-based* cognitive states that *reflect* the perspectival sensitivity of perception. A painter, for instance, might make a judgement about the apparent size of a tree while attempting to depict it. Other examples might include certain perceptual demonstrative thoughts, thoughts about particulars—standardly expressed through demonstratives, such as "this" or "here"—made available by one's perceptual relations to those particulars. Yet, the fact

that some perceptual feature, such as perspectival sensitivity, can be reflected in a perceptually-based cognitive state is no reason to hold that this feature is not a mark of the perceptual. After all, some features that are *clearly* perceptual, such as the apparent length of the lines in the Müller-Lyer illusion, can be reflected at the level of perceptual judgment. In the case of the Müller-Lyer, I may form the perceptually-based judgement that the lines look to be different lengths, even though I know they are not. The mere fact that I can form such a judgment does not show that the illusion has a non-perceptual basis. If anything, the fact that I cannot revise this perceptually-based judgement in light of my background knowledge is evidence that the illusion is a product of perception.

To assess the adequacy of the perspectival test, we should instead ask whether cognitive phenomena exhibit perspectival sensitivity *independently* of the perspectival sensitivity of perception. An affirmative answer is suggested by the fact that we often change our judgements and beliefs about things by literally changing our perspectival relations to them. For instance, I may believe that my bicycle is in the backyard but revise my belief if I change my perspectival relation to my bicycle by stepping outside. In turn, I may find that my bicycle was in my backyard *yesterday* but is now nowhere to be found. Notice, however, that this revised belief is not the result of a sensitivity to my perspectival relation to the bicycle *per se*. Rather, the change in my perspectival relation to the bicycle only results in a revised belief insofar as the change in relation yields new evidence. The change in perspectival relation plays only an instrumental role in changing my belief. I would not revise my belief if I were to step outside and find my bicycle in the same place as before. Likewise, I may have revised my belief in just the same way even if I did not step outside (e.g., if I were told that every bicycle in the neighborhood was stolen).

Similar considerations help to disarm other potential counterexamples to the perspectival test. Consider the game “hot and cold”, in which a small household object is hidden by player 1 while player 2’s eyes are closed. After the object is hidden, player 2 begins to search for it. As player 2 gets closer to the object, player 1 says, “hotter”. As player 2 gets farther away from it, player 1 says, “colder”. Now, player 2 may come to form the belief that they are *such-and-such* a distance away from the object, and this belief may be *evidentially* sensitive to the testimony of player 1, but the belief in question is not sensitive to player 1’s perspectival relations to the

object *per se*. Again, this is easy enough to appreciate: Had player 3 come along and secretly moved the object mid-game, player 2 would have still revised their belief in accordance with player 1's instructions.

5 THE PERSPECTIVAL SENSITIVITY OF RECOGNITION

In this section, I provide evidence that recognition exhibits perspectival sensitivity. In conjunction with the perspectival test, these considerations establish that recognition is at least partly perceptual.

By way of introduction, consider an experiment by Brady et al. (2008), in which subjects were instructed to remember 2,500 object images from a range of different object categories over the course of 5.5 hours. Each image was displayed once for three seconds. An additional task that required subjects to note image repetitions was also included so that subjects paid attention to each image. After this initial display, subjects were presented with two images, one "novel" and one "old" (i.e., an image selected randomly from the 2,500 object images displayed earlier). Subjects performed a task requiring them to select the image they had seen in the initial display of the object images. (This type of task is known as a "two-alternative forced choice" task.) In the "novel" condition, the novel image was of an object that differed in category from any of the objects depicted earlier. In the "exemplar" condition, the novel image was of a physically distinct object of the same category as the object depicted in the old image. In the "state" condition, the novel image was of the same object depicted in the old image but was displayed in a different state or pose. Changes in the "state" of an object are identity-preserving changes in its initial condition. For example, a briefcase undergoes a state change in this sense if it goes from being unopened to being opened. Similarly, a dresser would undergo a state change if one of its drawers were opened. Changes in the "pose" of an object are changes in the perspective or vantage point from which it was initially depicted. Across all three conditions, subjects were

remarkably accurate: They correctly identified the old image 93% of the time in the novel condition, 88% of the time in the exemplar condition, and 87% of the time in the state condition.

These results are interesting for many reasons.[13] The main point which I wish to extract from this study, however, are the results from the state condition. The results of this condition indicate that our capacity for recognition goes beyond the capacity to distinguish previously observed objects within or across categories. In the state condition, subjects were significantly above chance in distinguishing changes in the state or pose of a single object. The finding that subjects can distinguish changes in the pose of an object *would* provide evidence that in recognizing a particular subjects' responses vary across certain changes in their perspectival relations.

In Section 3, I claimed that registering that a particular (e.g., an object) has been observed before is necessary for recognition. Are participants in this experiment really registering that the objects depicted in the images in the forced choice task are the same objects depicted in the images they observed several hours earlier? Perhaps participants had a *raw feel* for which object image in the forced choice task was the correct one but were unable to *recognize* the depicted object. There is reason to think this worry is misplaced. As Brady and colleagues note: "Participant reports afterward indicated that they were usually explicitly aware of which item they had seen [several hours before], as they expressed confidence in their performance and volunteered information about the details that enabled them to pick the correct items" (2008, p. 14327).

Still, there is an alternative explanation of the participants' results. In the state condition, each trial involves either (i) a change in the initial condition of the depicted object (its "state"),

[13] For one, they provide a healthy contrast to change and inattention blindness studies. Such studies are frequently taken to show that we do not internally store rich worldly detail. For instance, Noë (2002) suggests that "work on change blindness seems to suggest that we may not in fact actually produce … detailed internal models" (p. 9). As Noë notes, change blindness studies concern unattended aspects of a scene. Granting that change blindness studies do not merely reflect difficulties with a post-perceptual comparison process, it may be reasonable to infer from these studies that we do not generate rich, internal models of those unattended aspects (cf., Simons & Rensink, 2005). By forcing subjects to allocate their attention appropriately, Brady and colleagues provide evidence that we retain richly detailed information concerning previously *attended* aspects of a scene. These results are likely even more striking in everyday contexts, where objects and scenes tend to be viewed multiple times and for durations extending beyond just a few seconds. In a task requiring subjects to answer questions concerning the details of previously viewed natural scene images, performance greatly improved as scene exposure times increased and when the scene was viewed multiple times before the questions were asked (Melcher, 2006).

(ii) a change in the vantage point from which the object was initially depicted (its "pose"), or (iii) a change in the object's state *and* pose. Given the inclusion of (i) and (iii), it is difficult to rule out the interpretation that subjects' strong performance in the state condition is primarily driven by their capacity to distinguish changes in an object's state, not its pose. If this were the case, one might argue that the study by Brady and colleagues does not provide strong evidence that recognition exhibits perspectival sensitivity. Instead, participants might only be sensitive to non-relational changes in the depicted object.

A discrimination task by Hollingworth and Henderson (2002) provides further evidence that recognition exhibits perspectival sensitivity while avoiding this latter worry. Using an eye tracker, the experimenters allowed subjects to freely view an image of a scene filled with various objects until they had directly fixated on a target object (e.g., a toy truck or a notepad). Once subjects had shifted their gaze to another part of the screen, the target object was covered with a speckled mask. After re-fixating the mask and pressing a button to indicate their readiness, subjects performed a two-alternative forced choice task in which they were to identify the original scene among two different versions displayed successively: The original version of the scene and a version of the scene that was indistinguishable from the original with the exception that the target object was rotated by 90°. (The order was counterbalanced across trials.) If subjects were *insensitive* to their perspectival relations to the target object in performing this recognitional task, their performance should be approximately at chance. Yet, participants were quite successful in recognizing changes in the orientation of the target object: They correctly identified the initial orientation of the target object well over 80% of the time. Similar results obtained in another experiment which involved much lengthier delays between the initial scene display and the discrimination task (i.e., 5–30 minutes) (2002, experiment 2).

One might worry that these studies do not show that subjects are sensitive to changes in their perspectival relations to particulars *per se.* Instead, subjects may only be sensitive to a particular's surface features whose visibility is affected by changes in perspectival relations. For example, certain changes to the orientation of a hammer might render more of its handle visible while simultaneously rendering less of its head visible. Maybe a sensitivity to differences in

these visible surface features alone explains the results of the above studies. If this is correct, these results do not establish the perspectival sensitivity of recognition.

A study by Standing et al. (1970) provides evidence in favor of the perspectival sensitivity of recognition while avoiding this worry. Using a two-alternative forced choice task, Standing and colleagues showed that subjects are capable of recognizing left–right mirror reversals in images. Even after significant 30-minute delays, subjects detected mirror reversals well over 80% of the time. These results are significant because mirror reversals alter one's perspectival relations to particular without affecting which of its surface features are visible. This indicates that the perspectival sensitivity of recognition cannot be "explained away" in terms of a mere sensitivity to differences in visible surface features.

Let me mention a final concern. The issue of whether recognition exhibits perspectival sensitivity is closely connected to debates concerning the effects of viewpoint on object recognition. Translated into the context of these debates, the foregoing results indicate that object recognition is sensitive to changes in an object's viewpoint (relative to the viewpoint from which it was initially observed). In general, however, the evidence that has emerged from these debates is mixed, and some studies suggest no effect of viewpoint on object recognition (see Hummel, 2013). Is this cause for alarm?

The answer would seem to be yes if the debate over the effects of viewpoint on object recognition is cast in "either/or" terms. Yet, this is probably a false dichotomy. Tarr and Hayward (2017) provide evidence that, by default, subjects concurrently encode *both* viewpoint-dependent and viewpoint-invariant information for the purposes of object recognition but flexibly produce viewpoint-dependent or viewpoint-invariant responses in object recognition tasks, depending on the nature of the task (see also Leek & Johnston, 2006).[14] These findings mesh with evidence that two neurally dissociable subsystems underlie object recognition—a "viewpoint-dependent" and a "viewpoint-invariant" system—whereby certain contextual factors lead to one system dominating the other (Burgund & Marsolek, 2000). If a story along these lines

[14] Translated into our larger discussion, these results suggest that object recognition exhibits *both* perspectival sensitivity and insensitivity—just as perception does.

is correct, we should *expect* studies to produce mixed results regarding the effect of viewpoint on object recognition.

6 RECOGNITION IS AN INTERFACE CAPACITY

The last two sections indicate that recognition is at least partly perceptual. However, the perspectival test used to establish this conclusion leaves open whether recognition is purely or only partly perceptual. In this section, I explore two lines of thought in support of the hypothesis that recognition is an interface capacity that is partly perceptual and partly cognitive. The first concerns the role of background information in recognition (Section 6.1). The second concerns the role of long-term memory (Section 6.2). To foreshadow, I find support from the second line of thought—with one qualification.

I advance our discussion by considering a model of recognition that has remained popular for several decades. I have in mind Bruce and Young's (1986) influential model of face recognition. In what follows, I omit details of the model that are not relevant for our purposes and focus on the model itself, rather than evidence in favor of it.

Let me flag two issues with casting the discussion in terms of Bruce and Young's model. First, the model is limited in scope and applies only to familiar face recognition. This restriction is not entirely artificial since recognition has been researched extensively in the context of faces. Moreover, it will become apparent that our conclusion straightforwardly generalizes beyond face recognition. Second, it is always possible that new findings will undermine the model. That said, the model has largely withstood the test of time. As Burton and colleagues remark: "In the 25 years, since Bruce and Young published their paper, the core theoretical distinctions it draws have survived remarkably intact" (2011, p. 953–954).

On Bruce and Young's model, face recognition begins with a process of "structural encoding" in which the visual system produces a structural code of a face. A "structural code" is

a representation of a face that abstracts away from certain transient features, such as hairstyle, while encoding certain invariant features of the face, such as the spatial arrangement of facial features. While some structural codes, such as those used for the analysis of facial speech, retain information concerning facial expression, those used for face recognition do not. They do, however, encode information concerning viewpoint.

These structural codes are transferred to "face recognition units" (FRUs), which store structural codes corresponding to previously observed faces—abstract long-term memory representations of previously observed faces. Each FRU yields an activation signal, which is a function of the degree of resemblance between the structural codes produced by the process of structural encoding and the structural codes stored in the FRU. A strong FRU activation signal is an indication that a face is familiar.

A FRU's baseline activation signal can also be increased by the activation of "person identity nodes" (PINs). Unlike FRUs, PINs are not modality-specific. For instance, a PIN can be directly activated by a voice recognition unit—the auditory analogs of FRUs. In turn, PINs can activate and be activated by other person-specific memories held in central cognition, such as memories about where or when some individual was born. It is through the activation of PINs that top-down influences are exerted on FRUs.

6.1 The role of background information in recognition

Let us turn to a first line of thought suggesting that recognition is an interface capacity. One might think that recognition cannot be purely perceptual because it requires a kind of decision, namely, a decision that the particular that is currently being observed is the same particular that has been previously observed. Since Helmholtz, it has been widely claimed that perception requires a (possibly figurative) form of decision-making or inference. For instance, it is often said that perceptual systems rely on an assumption that light comes from above to infer the three-dimensional shape of an object from two-dimensional patterns of light projected on the retina. Nevertheless, one might suspect that the kind of decision-making needed for recognition requires access to a sophisticated web of background information. One reason for thinking this is that

many of the particulars we observe and later recognize do not remain constant in appearance over time. And, if recognition is a task that requires unrestricted access to background information, one may worry that recognition is too computationally demanding a task for perceptual systems to solve without the aid of cognition.

On Bruce and Young's model, the kind of "decision" required for recognizing a face is rudimentary: A FRU's activation signal, *ceteris paribus*, depends on only: (i) the degree of resemblance between structural codes and (ii) the activation of certain PINs. Still, one might wonder whether such a model is too simple to account for the complexity of recognizing a face. We often see faces from distinct vantage points, under different lighting conditions, with varying expressions, and so forth. How is the model supposed to account for these complexities?

From the get-go, Bruce and Young were sensitive to these issues, speculating that structural codes encode only invariant features of faces that are stable over time. A similar approach is endorsed by Burton et al. (2005), who hypothesize that the stored representations that facilitate face recognition are produced by a process of averaging that preserves information concerning invariant facial features while eliminating transient information concerning transient facial features. If only invariant features are used for face recognition, superficial variations in facial appearance over time will not be disruptive. Longmore et al. (2008) offer a different approach, hypothesizing that face recognition is made possible by storing a collection of richly detailed representations, each produced as a result of a previous encounter with the face. As the collection corresponding to a particular face F grows, so does the likelihood that a novel representation of F will resemble one of the representations of F stored in the collection. Redfern and Benton (2019) suggest combining these two approaches.

Regardless, the systems responsible for face recognition account for variations in the appearance of a particular face over time *without* direct access to arbitrary background information. Face recognition is a complex task, but one that is completed using only face-specific information.

Consider the following rejoinder. The systems responsible for face recognition may not be informationally encapsulated from central cognition. After all, Bruce and Young's model

leaves open the possibility that central cognition can influence the activation signals of FRUs, at least indirectly through the activation of PINs. If one holds that perceptual systems are informationally encapsulated from central cognition, one may worry that face recognition cannot be purely perceptual.

Care must be taken to distinguish the well-established claim that perceptual systems are not *causally* encapsulated from central cognition from the contentious claim that perceptual systems are not *informationally* encapsulated from central cognition. As Wu stresses, in discussions of modularity of mind:

> The issue is not merely *causal* but *informational* encapsulation… reference to information is reference to semantic content over which computations are performed … establishing the failure of encapsulation [of system X from system Y] requires providing a mechanism where… computations [of system X] have access to and use [system] *Y* as an informational content resource (Wu, 2013, p. 656).

On Bruce and Young's model, FRUs cannot perform computations over information—that is, semantic content—stored in central cognition. This is because no information is transferred from central cognition to PINs to FRUs. For example, the information *Marcus is from Seattle* stored in central cognition is not transferred to PINs, even if the retrieval of this information activates a certain PIN. PINs do not store biographical information. Since interactions from central cognition to FRUs are mediated by PINs, the information is not transferred to FRUs either. The systems responsible for face recognition are thus informationally (but not causally) encapsulated from central cognition, according to Bruce and Young's model, because they cannot access information housed in central cognition, not even indirectly.[15]

[15] The distinction between causal and informational encapsulation adds a layer of interpretive complexity to studies which emphasize the role of conceptual processing in recognition. Schwartz and Yovel (2019) show that associating conceptual information with a face during an encoding phase facilitates later recognition. The experimenters rule out a range of explanations of this benefit (e.g., that it is due to more elaborate or global face processing) to show that it arises because of the formation of a conceptual representation. Nevertheless, the experimenters do not specify *how* this conceptual representation improves later recognition. On one interpretation, conceptually represented information improves recognition through an indirect causal influence on FRUs. On another interpretation, the

6.2 The role of long-term memory in recognition

Let us consider another line of thought suggesting that recognition is an interface capacity. Long-term memory representations play a constitutive role in recognition. On Bruce and Young's model, they take the form of structural codes stored within FRUs. These representations are "stimulus-independent" in the sense that they are not causally sustained by a distal stimulus by means of a present proximal stimulus, such as a pattern of light impinging on the retina (Beck, 2018). Yet, it is often said that, unlike cognitive representations, perceptual representations are stimulus-*dependent*, suggesting that the structural codes in Bruce and Young's model fall on the cognition side of the perception-cognition divide.

In assessing this line of thought, it is important to distinguish a FRU's activation signal from the structural codes stored within a FRU. In cases of successful face recognition, a face is a distal stimulus that causally sustains the activation signal of the appropriate FRU by means of a proximal stimulus. So, unlike structural codes, a FRU's activation signal is stimulus-dependent.[16] At the personal-level, these activation signals correspond to a stimulus-dependent recognitional response.

It might be argued that what distinguishes perceptual from cognitive representations is not stimulus-dependence *per se*, but that the former *function,* or aim*,* to represent in a stimulus-dependent fashion (cf., Phillips, 2017; Beck, 2018). One motivation for this amendment is that it allows certain representations to count as perceptual even in cases of hallucination; in such cases, a representation cannot be stimulus-dependent, though it might function to represent in a stimulus-dependent fashion. This amendment does not reduce the force of the present line of thought. It is not merely that long-term memory representations *are* stimulus-independent. It is apparent from reflecting on their role in our cognitive economy that they also function to represent in a fashion that is causally independent of any present (as opposed to past) stimulus.

---

systems responsible for face recognition perform computations over the conceptually represented information associated with the previously encoded face. Only the latter interpretation threatens the hypothesis that facial recognition systems are informationally encapsulation from central cognition. Yet, it is unclear why this interpretation ought to be favored over the former.

[16] To claim that a FRU's activation signal is stimulus-dependent is not to deny that there are other factors relevant to sustaining and altering it.

Thus, a long-term memory representation can be said to fulfil its function even if the represented stimulus no longer exists.

It might also be argued that certain perceptual representations nonetheless function to represent in a stimulus-independent fashion. The most compelling examples that come to mind involve so-called "assumptions" made by perceptual systems, such as the assumption made by the visual system that light tends to come from above. While new experiences may cause one to revise such assumptions (Adams et al., 2004), arguably, the aim of these assumptions is to represent the environment in a way that is not causally sustained by any present proximal stimulus.

Yet, a popular view is that these assumptions are not explicitly represented by perceptual systems at all. They are what Pylyshyn (1999) calls "natural constraints". As he puts it, "[t]he visual system does not need to access an explicit encoding of the constraint: it simply does what it is wired to do, which, as it happens, means that it works in accordance with the constraint discovered by the theorist" (p. 354; see also Orlandi, 2014). The visual system merely behaves *as if* it were operating under the assumption that light comes from above. If this is right, these assumptions cannot be examples of perceptual representations that function to represent in a stimulus-independent fashion.

Of course, a popular view is not necessarily a correct one. Although I am drawn to the natural constraints view, I shall not try to defend it here. Instead, I conclude with a qualification: Insofar as there are no perceptual representations that function to represent in a stimulus-independent fashion, long-term memory representations are cognitive representations. In turn, recognition is partly cognitive because it in part constitutively depends on such representations. Although our discussion has been cast in terms of Bruce and Young's face recognition model, this conclusion generalizes to any form of recognition that in part constitutively depends on long-term memory representations (e.g., familiar object or scene recognition).

7 CONCLUSION

I have argued that considerations relating to perspectival sensitivity indicate that recognition (in the token sense) is at least partly perceptual. Given the constitutive role of long-term memory representations in recognition, I found reason to think that recognition is also partly cognitive. In sum, recognition is an interface capacity that straddles the border between perception and cognition.

Our conclusion bears on the views set out by Strawson, Evans, and Peacocke (Section 1). One consequence of our discussion is that a creature incapable of forming long-term memories would lack a capacity for recognition. For Evans, this creature would lack a very basic means of identifying particulars non-descriptively; for Strawson, this creature would have *no* means of identifying particulars. On Peacocke's view, such a creature would be conceptually impoverished relative to us. It would not possess the concept "Lincoln Square", or any other recognitional concept for that matter. In short, if Strawson, Evans, and Peacocke are correct, long-term memory is significant for much more than remembering—it plays a foundational role in identification and concept possession as well.

Key questions remain. What is the role of affect in recognition? How are we to understand intellectually robust forms of recognition (e.g., recognizing a familiar mathematical proof)? What are the epistemic dimensions of recognition? These questions highlight that the topic of recognition is ripe for exploration. Unfortunately, we will have to wait until a later time to answer them.

ACKNOWLEDGEMENTS

I thank Ned Block, John Campbell, Dave Chalmers, Sam Clarke, Ravit Dotan, Eric Mandelbaum, Alva Noë, and two anonymous reviewers for comments and discussion on earlier drafts. This paper was also improved by the comments of participants at the annual meeting of the Society for Philosophy and Psychology in San Diego.

REFERENCES

Adams, W., Graf, E., & Ernst, M. (2004). Experience can change the "light-from-above" prior. *Nature Neuroscience, 7(10),* 1057–1058. doi: 10.1038/nn1312

Beck, J. (2018). Marking the perception–cognition boundary: the criterion of stimulus-dependence. *Australasian Journal of Philosophy*, *96(2),* 319-334. doi: 10.1080/00048402.2017.1329329

Begby, E. (2011). Review of Tyler Burge, *origins of objectivity*. *Notre Dame Philosophical Reviews, 2.*

Block, N. (2014). Seeing-as in the light of vision science. *Philosophy and Phenomenological Research, 89(3),* 560–572. doi: 10.1111/phpr.12135

Brady, T., Konkle, T., Alvarez, G., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105(38),* 14325-14329. doi: 10.1073/pnas.0803390105

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77(3),* 305–327.

Burge, T. (2010). *Origins of objectivity.* New York, NY: Oxford University Press.

Burgund, E., & Marsolek, C. (2000). Viewpoint-invariant and viewpoint-dependent object recognition in dissociable neural subsystems. *Psychonomic Bulletin & Review, 7,* 480-489.

Burton, A., Jenkins, R., Hancock, P., & White, D. (2005). Robust representations for face recognition: the power of averages. *Cognitive Psychology, 51(3)*, 256-284. doi: 10.1016/j.cogpsych.2005.06.003

Burton, A., Jenkins, R., & Schweinberger, S. (2011). Mental representations of familiar faces. *British Journal of Psychology, 102(4),* 943-958. doi: 10.1111/j.2044-8295.2011.02039.x

Byrne, A. (2005). Perception and conceptual content. In Steup, M., Turri, J., & Sosa, E., (Eds.), *Contemporary debates in epistemology* (pp. 231–250). Oxford: Wiley-Blackwell.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36(3),* 181-204. doi: 10.1017/S0140525X12000477

Drayson, Z. (2014). The personal/subpersonal distinction. *Philosophy Compass, 9(5),* 338-346. doi: 10.1111/phc3.12124

Evans, G. (1982). *Varieties of reference.* Oxford, UK: Oxford University Press.

Firestone, C., & Scholl, B. (2016). Cognition does not affect perception: evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *39*, 1–72. doi: 10.1017/S0140525X15000965

Flombaum, J., & Scholl, B. (2006). A temporal same-object advantage in the tunnel effect: facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance, 32(4),* 840-853. doi: 10.1037/0096-1523.32.4.840

Green, E. J., & Schellenberg, S. (2018). Spatial perception: the perspectival aspect of perception. *Philosophy Compass, 13(2),* 1-16. doi: 10.1111/phc3.12472

Green, E. J. Forthcoming. The perception-cognition border: a case for architectural division. *The Philosophical Review.*

Hollingworth, A., & Henderson, J. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28(1),* 113-136. doi: 10.1037/0096-1523.28.1.113

Huang, G., Ramachandran, S., Lee, T. S., & Olson, C. (2018). Neural correlate of visual familiarity in macaque area V2. *The Journal of Neuroscience, 38(42),* 8967– 8975. doi: 10.1523/JNEUROSCI.0664-18.2018

Hummel, J. (2013). Object recognition. In Reisburg, D. (Ed.), *Oxford handbook of cognitive psychology* (pp. 32–46). Oxford: Oxford University Press.

James, W. (1890). *The principles of psychology, volume I.* New York, NY: Holt.

Lande, K. (2018). The perspectival character of perception. *The Journal of Philosophy, 115(4),* 187–214. doi: 10.5840/jphil2018115413

Larzabal, C., Tramoni, E., Muratot, S., Thorpe, S., & Barbeau, E. (2018). Extremely long-term memory and familiarity after 12 years. *Cognition, 170*, 254-262. doi: 10.1016/j.cognition.2017.10.009

Leek, E., & Johnston, S. (2006). A polarity effect in misoriented object recognition: the role of polar features in the computation of orientation-invariant shape representations. *Visual Cognition, 13(5),* 573-600.

Longmore, C., Liu, C. H., & Young, A. (2008). Learning faces from photographs. *Journal of Experimental Psychology: Human Perception and Performance, 34(1),* 77-100. doi: 10.1037/0096-1523.34.1.77

Lupyan, G. (2015). Cognitive penetrability of perception in the age of prediction: predictive systems are penetrable systems. *Review of Philosophy and Psychology, 6*, 547–569. doi: 10.1007/s13164-015-0254-3

Mandelbaum, E. (2018). Seeing and conceptualizing: modularity and the shallow contents of perception. *Philosophy and Phenomenological Research*, *97(2),* 267-283. doi: 10.1111/phpr.12368

Mandler, G. (2008). Familiarity breeds attempts: a critical review of dual-process theories of recognition. *Perspectives on Psychological Science, 3(5),* 390-399. doi: 10.1111/j.1745-6924.2008.00087.x

Melcher, D. (2006). Accumulation and persistence of memory for natural scenes. *Journal of Vision, 6(2),* 8-17. doi: 10.1167/6.1.2

Morales, J., Bax, A., & Firestone, C. (2020). Sustained representation of perspectival shape. *Proceedings of the National Academy of Sciences, 117(26),* 14873-14882. doi: 10.1073/pnas.2000715117

Noë, A. (2002). Is the visual world a grand illusion? *Journal of Consciousness Studies, 9(5-6),* 1-12.

Noë, A. (2004). *Action in perception.* Cambridge, MA: MIT Press.

Orlandi, N. (2014). *The innocent eye: why vision is not a cognitive process.* New York, NY: Oxford University Press.

Peacocke, C. (1983). *Sense and content.* Oxford, UK: Oxford University Press.

Peacocke, C. (1992). *A study of concepts.* Cambridge, MA: MIT Press.

Phillips, B. (2017). The shifting border between perception and cognition. *Noûs, 53(2),* 316-346. doi: 10.1111/nous.12218

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22,* 341-423.

Quilty-Dunn, J. Forthcoming. Attention and encapsulation. *Mind & Language.*

Redfern, A., & Benton, C. (2019). Representation of facial identity includes expression variability. *Vision Research, 157,* 123-131. doi: 10.1016/j.visres.2018.05.004

Ricker, T., & Cowan, N. (2010). Loss of visual working memory within seconds: the combined use of refreshable and non-refreshable features. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36(6),* 1355-1368. doi: 10.1037/a0020356

Rosch, E. (1978). Principles of categorization. In Rosch, E., & Lloyd, B. (Eds.), *Cognition and categorization* (pp. 27–48). Oxford: Lawrence Erlbaum.

Russell, B. (1921). *The analysis of mind.* London, UK: Allen and Unwin.

Schellenberg, S. (2008). The situation-dependency of perception. *The Journal of Philosophy, 105(2),* 55-84.

Schwartz, L., & Yovel, G. (2019). Learning faces as concepts rather than percepts improves face recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45(10),* 1733-1747. doi: 10.1037/xlm0000673

Schwitzgebel, E. (2006). Do things look flat? *Philosophy and Phenomenological Research, 72(3),* 589-599. doi: 10.1111/j.1933-1592.2006.tb00585.x

Simons, D., & Rensink, R. (2005). Change blindness: past, present, and future. *Trends in Cognitive Science, 9(1),* 16-20. doi: 10.1016/j.tics.2004.11.006

Standing, L., Conezio, J., & Haber, R. (1970). Perception and memory for pictures: single-trial learning of 2500 visual stimuli. *Psychonomic Science, 19(2),* 73-74.

Strawson, P. F. (1959). *Individuals.* London, UK: Methuen.

Tarr, M., & Hayward, W. (2017). The concurrent encoding of viewpoint-invariant and viewpoint dependent information in visual object recognition. *Visual Cognition, 25(1-3),* 100-121. doi: 10.1080/13506285.2017.1324933

Tye, M. (2000). *Consciousness, color, and content.* Cambridge, MA: MIT Press.

Wagner, M. (2006). *The geometries of visual space.* Mahwah, NJ: Lawrence Erlbaum.

Wu, W. (2013). Visual spatial constancy and modularity: does intention penetrate vision? *Philosophical Studies, 165,* 647-669. doi: 10.1007/s11098-012-9971-y

Yonelinas, A. (2010). Recollection and familiarity: examining controversial assumptions and new directions. *Hippocampus, 20(11),* 1178–1194. doi: 10.1002/hipo.20864

Zahorik, P., & Wightman, F. (2001). Loudness constancy with varying sound source distance. *Nature Neuroscience, 4(1),* 78-83. doi: 10.1038/82931

Zhang, W., & Luck, S. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science, 20(4),* 423-428. doi: 10.1111/j.1467-9280.2009.02322.x