The Varieties of Categorization and the Perception-Cognition Debate

Henry Shevlin

Leverhulme Centre for the Future of Intelligence, University of Cambridge

Correspondence: Hfs35@cam.ac.uk

The Varieties of Categorization and the Perception-Cognition Debate

Our experience of the world typically makes us aware of a wide variety of properties of objects, ranging from their color and shape to their identity and function. In this paper, I use the notion of categorization to explore these various forms of awareness. I begin by arguing for a tripartite model of categorization that loosely groups our awareness of the properties of objects into three broad classes, namely low-, mid-, and high-level categorization. I relate this tripartite distinction to the current controversies surrounding the perception-cognition distinction, and argue that, of the three classes of categorization, mid-level categorization is of interest. Roughly, this is the process by which we effortlessly and automatically recognize perceived objects as being of given kinds. I go on to consider the question of whether mid-level categorization is best considered a form of perception or cognition, and examine merits and difficulties associated with either approach. Finally, I explore a more radical approach I term the Three Kinds view that takes mid-level categorization to be best understood as a further process intermediate between perception and cognition.

Keywords: perception, cognition, memory, categorization

1. Introduction

In experiencing the world, we come to be aware of the properties and identities of objects and scenes around us. Some of these properties are simple attributes like color, shape, and motion, but we also quickly learn to recognize specific individuals, different types of animal, and various kinds of artefacts. However, it is by no means introspectively or pretheoretically obvious how we should locate these various abilities within the architecture of the mind. There is little consensus, for example, concerning whether our ability to recognize specific people is a function of the same core perceptual mechanisms as our ability to recognize color and shape, or whether it depends on a further psychological process.

In this paper, I offer an overview of this very broad suite of abilities, which I will loosely refer to as *categorization*, and draw out some relevant contrasts between the various ways in which we become aware of the world. Needless to say, of course, this project is hardly novel in substance: categorization in this sense has been the focus of extensive philosophical and scientific investigation. However, such investigation often proceeds by asking more specific questions about, for example, the contents of visual experience (Siegel, 2010), the role of concepts in perceptual awareness (McDowell, 1994), the relationship between perceiving and perceiving-as (Dretske, 2004), or the role of attention in the extraction of semantic information (Broadbent, 1958). By approaching the question from a relatively theoretically neutral position and examining some of the different ways we categorize the world, my hope is that we might glean

new insights into the different processes underlying various forms of awareness and how those processes are structured within the mind.

The paper proceeds as follows. In section 2, I survey some differences in the ways that categorization in the broadest sense makes us aware of properties of objects in the world, and develop a broad distinction between what I term low-, mid-, and high-level forms of categorization. In section 3, I consider how this distinction maps on to models of the mind that draw a sharp contrast between perception and cognition. In section 4, I examine more closely the question of where to locate mid-level categorization in such frameworks. Finally, in Section 5, I consider a more radical proposal, the 'Three Kinds' view, that claims that mid-level categorization might be best considered as a third capacity distinct from both perception and cognition.

2. The Varieties of Categorization

2.1. A tripartite schema for categorization

My first task, then, will be to spell out in a little more detail what I mean by categorization. Construed very broadly, I suggest categorization can be understood as the representing of objects, individuals, and events in respect of some repeatable attributive element; that is, to represent something as *being F*. According to this (very expansive) way of understanding categorization, seeing an object as blue, hearing a noise as a siren, thinking of my friend Anita as a generous person, or imagining the future as bleak can all be considered instances of categorization. Indeed, one might wonder whether there are *any* mental states that all together lack a categorical element.[1]

Nonetheless, we can gain purchase on the various ways in which categorization occurs in the mind by noting that in everyday experience, we come to be aware of the properties of the world around us in seemingly different ways. One such form of awareness involves the rapid perceptual detection of the basic qualitative properties of objects around us – color, size, shape, motion, and so on. Thus, as Prinz puts it, "[i]n vision, we experience stimuli as bounded wholes from a specific vantage point, occupying a specific size and position within the visual field" (Prinz, 2011). To give a mundane example, someone waking up in the morning with bleary eyes might initially be aware of the objects around them primarily in respect of color, shape, and size, before recognizing them as, say, a light fixture, a window, and a cabinet. The ability to perceive objects in respect of such low-level properties is (setting aside pathological cases) common to most creatures with the requisite sensory capacities, and need not rely on specific learning histories. Putting this another way, every neurotypical sighted human who has ever lived has been able to see colors, shapes, and motion.

---

[1] Dretske suggests that perception can occur without categorization. He claims, for example, that "I do not have to notice, recognize, or identify every book I see on the shelves as I scan them in search of a particular book" (2004: 10). Similarly, many Disjunctivists would likely deny that veridical perception need involve the representation of objects as exemplifying some property or attribute.

However, in addition to this set of basic perceptible properties, we are frequently aware of objects in respect of classifications learned only through experience. In order to see a seagull *as* a seagull, for example, I must – in some sense to be spelled out – have undergone a suitable number of prior experiences of seagulls, or perhaps to have learned to distinguish them from other birds, or to have come to understand that they form a distinct kind, and so on.

Setting aside the precise mechanisms via which this awareness is achieved – whether via association, conceptualization, or some dedicated form of perceptual learning – note that such awareness of the identities and properties of objects is ubiquitous and important. Our ability to navigate the world relies upon our immediate and effortless recognition, without pause for reflection or conscious inference, of this object being a car, that object being a £10 note, that person as being our friend David, and so on. I term this capacity – to be spelled out in more detail below – *mid-level* categorization. This contrasts with what I will term the kinds of *low-level* categorization described above, which are constrained in their representational repertoire to certain very general attributes.[2]

There are of course further ways in which we come to be aware of the categorical identity of objects, notably via a process of inference from has already been perceived. Thus, as I look at various animals at the zoo, I may see a duckbilled platypus, and consequently judge that it is a monotreme. Such forms of inference on the basis of what we see are similarly ubiquitous in daily life. Possible examples may include, for example, recognizing that a car is a particularly unusual model, noticing that a colleague is being evasive in her answers, seeing that a text is in an unfamiliar language, or reading off the precise time from an analogue clock. What I take to be common to these cases, at least typically, is that the relevant process of categorization follows our initial awareness of a scene with some (perhaps slight) delay. We can thus distinguish this *high-level categorization* from the forms described in the preceding paragraphs.[3]

At a first pass, then, we can distinguish three kinds of categorization that accompany experience. What I term low-level categorization involves immediate perceptual awareness of basic properties such as shape, color, and so on. Intermediate-

---

[2] One might reasonably deny that mere innate sensitivity to basic perceptible properties is properly called categorisation at all (Dretske, ibid.). While I will continue to refer to this capacity as low-level categorisation, I do not take any of my arguments to rest on the point; if someone were to insist, for example, that the term "categorisation" should be reserved for the processes I term mid- and high-level categorisation, it would have little substantial impact on the arguments that follow.

[3] Though appeals to the kinds of properties represented in experience are helpful in fleshing out the distinction between the various forms of categorisation, I do not take the distinction to fundamentally rest on this point. Thus, looking down the road on a dark night, I might judge after a moment's delay on the basis of gait that a figure in front of me is my friend Anita. Insofar as this involves an inference from initial perceptual awareness, it would count as a case of high-level categorization. Similarly, one might imagine cases in which relevant experts are immediately aware of, say, a given animal as a monotreme, or a car as a 1961 Ferrari 250GT California, where the immediacy of the awareness makes these instances of mid-level categorization.

level categorization is similarly subjectively immediate, but involves awareness of learned kinds and classifications. Finally, high-level categorization involves a slower and more clearly inferential process of coming to believe that some object one has already perceived falls under a specific category.

2.2. Empirical support for the tripartite schema

I take this way of carving things up to have some pretheoretical appeal. However, the distinction as described thus far is somewhat rough-and-ready. There are thus borderline cases whose place within the framework may be initially unclear: we might well ask, for example, whether our tactile awareness of objects as heavy is properly considered a low- or intermediate form of categorization, or whether our recognition of someone's accent as German is more akin to mid- or high-level categorization. Far more is needed if we are to defend this schema as having any claim to being explanatorily useful for cognitive science or introspective psychology.

With this in mind, I would suggest that this tripartite division among forms of categorization meshes well with important findings from empirical science. Consider first the distinction between mid- and high-level categorization. One experiment by Grill-Spector and Kanwisher (2005) arguably shows how this distinction might be spelled out scientifically. Grill-Spector and Kanwisher presented items to subjects for brief intervals (ranging from 17-200ms) and examined their ability to 'detect', 'categorize', and 'identify' them. The detection task required subjects to decide whether an object (as opposed to a texture) had been presented. The categorization task used a similar methodology but required subjects to press a button to indicate whether an item from a target category (e.g., car vs. not-car) had been displayed. Finally, the identification task required subjects to indicate whether the item belonged to a within-category class (e.g., German Shepherd vs. some other dog). Grill-Spector and Kanwisher discovered that while subjects were slower on the fine-grained identification task, their accuracy and speed was just as fast for the categorization task as it was for the detection task. Summarizing their results, they note that "detection and categorization performance require the same amount of information and processing time" and that "[b]y the time subjects knew an image contained an object at all, they already knew its category."

This result supports the idea that there is an important distinction to be drawn between the kind of awareness of basic categories that immediately accompanies perception and the more fine-grained discriminations that follow it. In other words, it suggests that some categorical information does not rely on making inferences that occur only after initial perceptual experience. Note also that subjects' relative slowness in fine-grained identification tasks as compared to basic categorization tasks supports the idea that the latter typically involves the kind of relatively generic semantic categories ("dog", "guitar", "car", and so on), which we might intuitively associate with mid-level categorization.[4]

---

[4] However, see preceding footnote.

Consider next the distinction between low-level and mid-level categorization. Grill-Spector and Kanwisher's experiment suggests that we are not typically aware of low- and mid-level categories at different stages of perceptual experience; otherwise, we might expect subjects to be faster at detecting the presence of generic objects than recognizing their category. Despite this, there is some neurological evidence which suggests that this awareness relies on different stages of processing. Specifically, patients with associative agnosia can perceive (and accurately copy) images in respect of color, size, and shape, yet are unable to recognize what they are seeing. This suggests that there is more to seeing something as a given object than simply perceiving its shape, size, color, and so on. By contrast, a condition known as *apperceptive* agnosia leaves patients unable to identify even low-level properties such as color, shape, and motion (Duffy, 1999). What these two conditions arguably suggests is that the underlying mechanisms for mid-level are somewhat distinct from those involved in awareness of color, shape, size, and other features I take to be low-level.

A third important source of evidence for exploring the proposed tripartite distinction concerns the impressive performance of non-human animals in many categorization tasks. For example, note that pigeons can be trained to discriminate novel pictures based on whether they contain images of people, and also seemingly learn to discriminate novel beach scenes from scenes of mountains or streets (Kirkpatrick et al., 2014). Rats, too, can learn to make same-different discriminations that can be extended to novel stimuli, and can learn to distinguish novel pictures of chairs, flowers, cars, and humans on seemingly categorical grounds (Brooks et al. 2013). Perhaps most strikingly of all, bees can be trained to make same/different and above/below distinctions for wholly novel stimuli (see Chittka & Jensen, 2011 for a review).

In light of the preceding discussion, the question arises whether we should think of these abilities as involving low-, mid-, or high-level forms of categorization. Consider first the idea that they should be understood in terms of low-level categorization. Insofar as these examples involve animals learning, in effect, to develop general recognition schemata that can be applied to novel stimuli, it is not clear how a purely low-level account could accommodate them. Recall that, by definition, the representational repertoire of low-level categorization is confined to very general perceptual attributes like color, shape, size, and so on. While such qualities are vital to determining the categorical identity of a given stimulus, the mere ability to represent something as, say, red and square on a given occasion does not automatically suffice for being able to form a new categorization class that can be applied to novel stimuli.

Nonetheless, perhaps we can understand these cases as involving complex 'shape gestalts' or generalized sensory templates (Siegel, 2010:111-12). For example, when a pigeon learns to identify beach scenes, we might explain this in terms of its visual system developing an abstract geometric and color template that is activated when beach scenes are presented. Certainly, it seems likely that *some* such process is at work when we learn to visually identify a new category: in learning to categorize a new kind of object or scene, such abstractions from low-level sensory information are all that we may have to work with.

I would question, however, whether such an account is best described as a low-

level view at all. That is to say, once we grant that some further psychological process – such as generalization or abstraction on the basis of low-level information – is required to reidentify stimuli as belonging to a given category, we are positing an additional mechanism beyond that involved in simple cases of low-level categorization. At this point, I am inclined to think that we have simply arrived at a particular way of articulating what is involved in mid-level categorization.

Moving on to an alternative view, we might wonder whether, in at least some of these cases, animals are engaged a form of high-level categorization. The idea, roughly, would be that animals perform some kind of post-perceptual inference that allows them to identify the object or scene in question. A challenge for this view is that it is far from clear whether we can reasonably assert that animals possess the kind of inferential capabilities characteristic of high-level categorization. Certainly, it seems doubtful that simpler creatures will possess a capacity for generalized propositional inference (however, see Carruthers, 2009).

A less radical alternative might be to suggest that the relevant mechanism is merely an associative one. One might worry, however, whether mere brute association should qualify as a case of categorization at all. Imagine, for example, that I hear a song which I have previously heard only in highly stressful contexts, and consequently feel a rush of fear. In such a case, it does not seem right to say that I *categorize* the song as fearful; rather, the song serves as a relevant trigger for the emotion.

It may nonetheless be the case that some apparent cases of animal categorization could be explained away as, in fact, instances of association. Whether this is correct – and which putative examples of animal categorization it would apply to – is of course a largely empirical question. However, insofar as there are at least some instances in which animals can genuinely categorize objects in their environment in respect of some learned attributive, I suggest that these abilities are best understood as a form of mid-level categorization. If this view is correct, it in turn lends some further reason to think that the tripartite distinction as described above might allow us to draw useful distinctions across different kinds of recognitional capacities.

3. Categorization and the Perception-Cognition Debate

3.1 – The perception-cognition controversy

The above examples constitute a very small sample of the relevant empirical data on the various forms of categorization. However, I hope that they help to sharpen and make somewhat more plausible the distinction between the various stages of categorization that accompany awareness. Additionally, they may serve to illustrate how the proposed tripartite division across different forms of categorization might be fruitfully put to work in empirical science and thereby earn its explanatory keep.

With this distinction in place, then, I turn to the central task of the remainder of the paper, namely understanding how these forms of categorization relate to different parts of the mind, and in particular to perception and cognition. Which, if any, of the forms of categorization described above are proper functions of

perception, and which functions of cognition?

Before proceeding, note that the very existence of a perception-cognition distinction is a matter of controversy, with debates dating back at least to Plato's *Republic* (602c-e). In more recent years, the idea that perception and cognition have quite different features and are underpinned by distinct mechanisms has been notably championed by philosophers such as Dretske (1981), Fodor (1983), Burge (2010), and Block (2014). It has also been challenged by Bayesian and predictive coding accounts of the mind championed by theorists such as Gary Lupyan (2015) and Andy Clark (2013). Additionally, the growing philosophical literature concerning alleged 'top-down' effects on perception from cognition has threaten to blur the boundary between the two processes.[5]

Nonetheless, in what follows I will operate on the assumption that there is an intelligible distinction between perception and cognition to be had. This is because the question of where to locate the various forms of categorization in the mind may seem uninteresting or even ill-formed for theorists who reject a perception-cognition distinction outright. If someone holds, for example, that the functions commonly associated with perception and cognition are in fact best understood as continuous processes involving multiple layers in a hierarchical network (thereby perhaps making the perception-cognition distinction otiose), they might naturally deny that there is any real question about which *specific* part of mind accomplishes the different forms of categorization.

Hence while I consider such ideas to be live options in current debates about the mind, what follows will be primarily applicable to theories that do accept the existence of a distinction between perception and cognition, and will explore how they can best accommodate the forms of categorization listed above.

3.2 – Fleshing out the perception-cognition distinction

Even among theorists who accept that there is some robust distinction between perception and cognition, it is nonetheless far from clear precisely how it should be drawn (Block, 2014; Burge, 2010; Carey 2009). Prima facie, at least, there are *many* important differences between the roles of perception and cognition. Before proceeding, then, it is worth briefly fleshing some of these out.

Note first that perception and cognition seem to exhibit a number of simple functional differences. Perception operates rapidly and automatically, whereas many cognitive processes such as reflective reasoning typically operate slowly, and are at least potentially under voluntary control. Likewise, perception and cognition seem to differ in their representational repertoire. Even if one claims, for example, that we can perceive natural kinds or causal interactions (Siegel, 2010; Bayne, 2009), we surely do not perceptually represent highly abstract theoretical kinds like Gini coefficients or haloclines, nor complex logical structures like conditionals. By contrast, we have no

---

[5] However, note that the mere existence of top-down effects need not count as decisive evidence against the idea that there is a robust perception-cognition distinction (see Block, 2016).

difficulties in thinking or forming desires about these things. Conversely, it has sometimes been suggested (by, e.g., Peacocke, 1992) that perception has a finer grain than cognition, allowing us to perceive far more shades of red, for example, than we can form thoughts about.

Such differences between perception and cognition are often connected with the claim that they possess distinct representational formats. Dretske (1981) and Tye (1995), for example, take it to be partly constitutive of perception that its content has an *analogue* format, unlike the 'digital' or discrete contents of thought. Similarly, Block (2014) and Burge (2010) claim that perception involves nonconceptual content and a sensory or iconic format, whereas thought is conceptual and propositional.

Perceptual psychology has also revealed less immediately obvious differences between perception and cognition. For example, the short-term memory mechanisms often thought to underpin perception have a large capacity (Sperling, 1960; Block, 2007), whereas our ability to perform many cognitive tasks seems to be limited by the austere capacity of working memory (Cowan, 2001). Additionally, it is sometimes suggested that perception can be distinguished from cognition on the grounds that the former is *encapsulated* from the latter (Fodor, 1983). That is to say, the contents of perception may not be amenable to correction or direct modulation by cognitive processes. Hence if I am looking at an image like the famous Muller-Lyer illusion, I cannot bring myself to *see* the lines as equal in length, even though I know they are.

Finally, there is arguably some phenomenological appeal to the perception-cognition distinction. Perceptual and other sensory forms of experience are famously characterized by their qualia, the phenomenal characters associated with experience of qualities such as redness, pain, and saltiness. By contrast, it remains controversial whether canonical forms of cognition such as understanding and reflective thought have any phenomenal character at all; even those sympathetic to the idea that there is a phenomenology of thought must allow that its subjective character seems somewhat different from seeing and hearing, for example.

Bringing these considerations together, we might sketch a very broad position that I will term the Distinct Kinds view. In essence, this takes it to be the case that perception and cognition differ constitutively in respect of one or more attributes (see Fig. 1 below). While theorists are likely to disagree about precisely which attributes are essential and which incidental to the perception-cognition distinction, what I take to be common among them is the idea that perception and thought differ in psychological role, phenomenology, content, or some combination of the above, and are thus quite reasonably considered distinct processes.

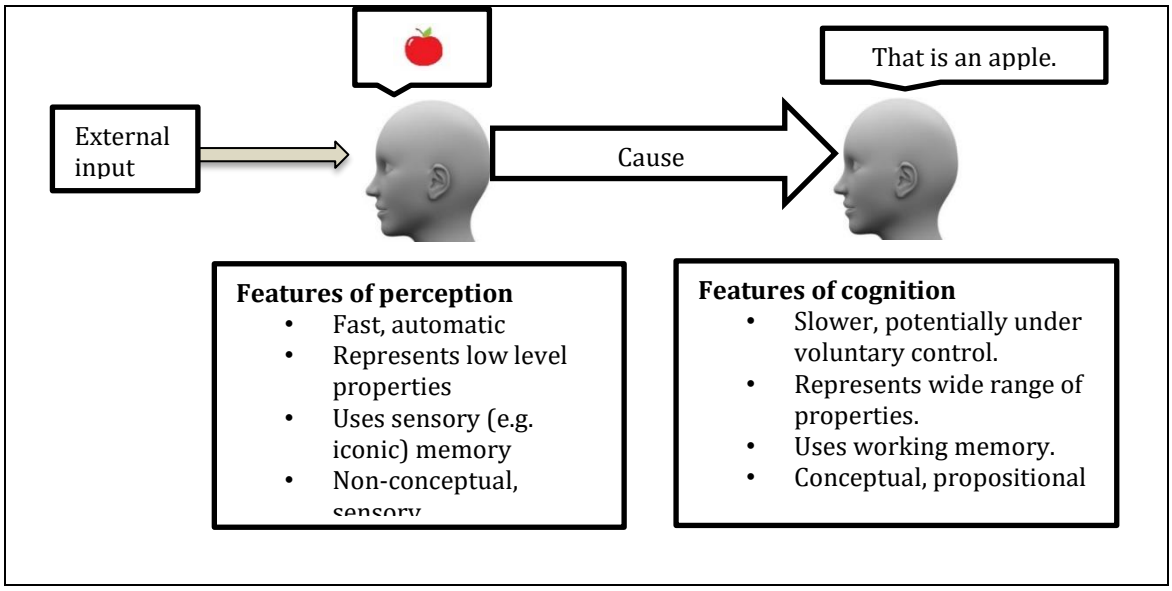

Fig. 1. A schematic illustration of the Distinct Kinds view

3.3 – Categorization and the distinct kinds view

Using the Distinct Kinds view as a guide, we can now ask where the various forms of categorization describe in Section 2 fall on the perception-cognition divide. I take it that, if any form of categorization is robustly perceptual, it is *low-level* categorization. Low-level categorization as characterized above is fast and automatic and represents low-level properties. By the lights of many theorists (notably Block, 2007), it relies on sensory memory and possesses nonconceptual contents. Indeed, if low-level categorization is not perceptual, it is not clear what else could count as perception.

High-level categorization, by contrast, seems to fall squarely on the side of cognition. As characterized earlier, high-level cognition is by definition slower than low- and mid-level categorization, and has an expansive repertoire of contents. While it may not always be under voluntary control (we sometimes make involuntary inferences, for example), it seems that it can in at least some cases be brought under such control; one might look at an animal and refrain from inferring its biological family, for example. Finally, insofar as we take high-level cognition to be characteristically inferential and not merely associative (as briefly discussed in 2.2, above), it lends itself to being fleshed out in terms of propositionally-structured inferences with conceptual contents.

Setting low- and high-level cognition aside, then, this leaves only the status of mid-level categorization unclear. While mid-level categorization shares some of the features frequently associated with perception (in being fast and automatic, for example), it also seemingly involves acquired high-level contents (like ‘Anita’ or ‘owl’) of a kind characteristic of canonically cognitive processes such as judgment, inference, and belief.

One radical possibility, to be explored in Section 5, is that mid-level categorization might best be viewed as distinct from both perception and cognition as a faculty in its own right. However, I regard it as an open possibility that mid-level

categorization might be accommodated within existing accounts of perception or of cognition, or some closely related capacity such as perceptual imagination. In the next section, then, I will briefly survey the merits of these approaches.

4. Is Mid-Level Categorization Perceptual or Cognitive?

4.1. Mid-level categorization as high-level perception

Let us first examine the idea that mid-level categorization should be considered part of perception proper. Broadly speaking, this would involve a commitment to the idea that part of the function of perception is to represent – in addition to color, shape, size, and so on –attributes such as dog, car, fire, and so on.

There are a variety of ways such views can be elaborated, with a key point of contention concerning the nature of content involved in such mid-level categorization. On one approach, both low- and mid-level categorization would be functions of perception proper, and both would involve states with non-conceptual content (this may be the view of Siegel, 2010). A different sort of account one might claim that perception is purely conceptual or has a conceptual component, perhaps with concepts 'bound in' to a relevant non-conceptual representation (Carruthers, 2014; Fodor, 2015). In that case, one might maintain that low- and mid-level categorization, though both perceptual, differ in respect of the kinds of content involved.

I regard both kinds of approach as viable and worthy of further elaboration. A feature common to them, however, is that they arguably require us to understand perception as a complex faculty involving distinct kinds of representational capacity. In the case of the non-conceptual approach, note first the view is seemingly committed to there being non-conceptual representations with learned categorical content like 'pine tree' or 'Barack Obama'. Unlike low-level categorization, these kinds of contents seem to involve representations with a discrete or 'digital' format (Dretske, 1981; Tye, 1995). That is to say, whereas an object's color, shape, size, and so on are arguably continuous qualities that come in degrees, being a car or being Barack Obama seems to be an all-or-nothing matter. [6]

It is natural to interpret this difference in terms of a further difference in representational format: whereas color, shape, motion, and so on possess an *iconic* representational format, discrete categories such as 'Barack Obama' have a *discursive* format (Fodor, 2007; Quilty-Dunn, 2016). Putting this in less technical terms, most if not all of the outputs of low-level categorization seem to have a syntactic structure that allows for representations of continuous variation. By contrast, the contents

---

[6] Note that this claim does not commit one to the view that mid-level categorization lacks degrees of determinacy (Stazicker, 2018) or confidence. It is possible that a given stimulus may be categorized as 'dog', for example, in such a way that many of its features are unspecified, or with a lesser or greater degree of subjective confidence. However, this does not undermine the point that the category *dog* either is or is not tokened in a given instance.

characteristic of mid-level categorization are such that they simply apply or fail to attribute a given category to an object. This significant difference in content then prompts the question of how a single system could generate representations of such different kinds, and arguably the most appealing answer involves appeal to a difference in the relevant *vehicles* of representation. And if this is correct, then the nonconceptual theorist might do well to allow that perception involves quite different kinds of representations.

A similar point can be made in response to the conceptualist approach. Insofar as we grant that perception involves both analogue sensory contents like color and discrete conceptual contents like 'pine tree', we are left with a difference in representational content that is arguably mysterious without appeal to differences in the relevant representational vehicles. One option here, of course, would be to deny that even low-level categorization involves nonconceptual contents (e.g., McDowell 1994). Note, however, that this approach has far more plausibility as an account of the contents of perceptual *experience* rather than perception per se. It is far harder, in light of the wealth of data from perceptual psychology, to insist that perception, as opposed to just perceptual experience, involves exclusively discrete representations, especially in the early processing of sensory information (Beck 2015). Hence while this approach may avoid a duality in the content and vehicle of representations at the level of conscious experience, it will struggle to escape such a duality at the level of perception *tout court*.

To summarize, then, I would suggest that perceptual approaches to mid-level categorization are on their strongest ground when they grant that the representational mechanisms underpinning low- and mid-level categorization are distinct subprocesses within perception, quite possibly involving different sorts of representational vehicles.[7] However, this might prompt the question of why we should consider both kinds of process to be underpinned by a single psychological faculty, namely perception, rather than the two processes being distinct faculties for categorizing and representing the world.

Of course, there are entirely reasonable answers that could be given to this question, and the idea that perception involves two distinct forms of representation is likely to be independently attractive to some theorists. Nonetheless, in adopting this dualistic view of perception, we do forego the idea – which has been found appealing by many – that perception is individuated at least in part by its possessing contents and vehicles of single underlying kinds.

4.2. Mid-level categorization as low-level cognition

A different strategy from the perceptual accounts sketched above would be to claim that mid-level categorization is in fact a form of cognition proper. There are a wide

---

[7] This approach also has the advantage of suggesting an immediate account of associative agnosia: once we allow that low- and mid-level categorisation involve at least somewhat distinct psychological processes, we can readily imagine how the latter could be impaired without disruption of the former.

variety of ways that one may spell this out, depending on one's pre-existing commitments to the constitutive characteristics of cognitive (as opposed to perceptual) processes. For example, one might identify some forms of mid-level categorization with rapid associative thought, or cognitive maps (Rescorla, 2009), or even imagination (Prinz, 2011).

However, I would suggest that such approaches face a similar consequence to that raised above in relation to perceptual accounts: while there are appealing ways one might understand mid-level categorization in terms of cognition, this will likely involve granting that cognition involves at least two quite different representational processes. More specifically, I suggest that insofar as mid-level categorization arguably lacks some of the features associated with canonical forms of cognition like belief, judgment, and inference, any account that incorporates it within a general faculty of cognition will do so only insofar as it is willing to grant that cognition is not a monolithic psychological system, but involves quite different representational processes.

To illustrate this, it will be helpful to examine a simple cognitive account of mid-level categorization and identify the challenges it faces. Consider the view that mid-level categorization is best understood as the forming of rapid propositional *judgment* or *beliefs* (Byrne, 2009; see also Pitcher, 1971). In other words, when I perceive something as a pine tree or Barack Obama, I am making a kind of post-perceptual judgment.

Such post-perceptual judgments doubtless do occur in many cases (indeed, I take them to be characteristic of high-level categorization), but insofar as we wish to claim that they can explain mid-level categorization, we must allow that they differ markedly from canonical cases of judgment and belief, such as judging that it will rain tomorrow. For one, many if not all cases of mid-level categorization are not readily amenable to correction in light of my background beliefs. This as illustrated by pareidolic illusions, like the images shown below. Even if we are informed and confidently believe that what we are seeing, for example, is a pool of soap bubbles, we cannot help but see it *as* an eye.[8]

[8] Note that unlike simpler illusory cases like the Muller-Lyer illusion, pareidolia need not involve any misrepresentation of low-level properties like colour or size, but only the categorical identity of what is seen. Pareidolia, then, constitutes a distinctive kind of *categorical* known illusion case.

Fig. 2. Examples of Pareidolia: an eye in soap bubbles and an owl in a coffee cup.

This illustrates a difference, then, between mid-level categorization and everyday instances of judgment: whereas our judgments are typically sensitive (to some degree) to other information at our disposal, no matter how firmly we believe that what we see is simply bubbles in a sink, we are seemingly irresistibly led to see the image as depicting an eye.

This does not by itself illustrate that mid-level categorization is of a different psychological species than judgment; there are, after all, some cases where judgments are seemingly insensitive to correction in light of new information. However, it should be noted that it would be quite unlike the familiar roster of cases where our beliefs and judgments are not amenable to fast correction. Biases, motivated reasoning, and simple forgetfulness, for example, might lead someone to hold contradictory beliefs at different times or fail to update their judgments in light of new information. Pareidolia, however, presents a more difficult case, insofar as I may *simultaneously* have two occurrent representations of the image above with contradictory contents, judging "that is not an eye" while continuing to see it as an eye. There is, to my knowledge, no clear example of such simultaneous contradictory contents occurring in uncontroversially cognitive processes. Similarly, whereas most of our beliefs are at least in principle open to correction, albeit in some cases with difficulty, it is hard to imagine what kind of purely *cognitive* shift could ever lead someone to stop seeing the image on the left as an eye.

Whatever kind of cognitive process mid-level categorization may be, then, it must differ in at least this respect from canonical judgment and belief. Another important difference concerns the contrast in phenomenology between mid-level categorization and everyday judgment. Compare the case in which I see an animal as a kangaroo and a case in which, on seeing a kangaroo, I form a rapid judgment to the effect that the animal in front of me is a marsupial. Phenomenologically, I would suggest, in the former but not the latter case the relevant category is somehow instantiated in my perceptual awareness: I am seeing the kangaroo as a kangaroo, but not strictly *seeing* it as a marsupial. This is not necessarily to claim that there is any phenomenology associated with being a kangaroo per se, but rather, that my awareness of the animal as a kangaroo is integrated with my perceptual awareness in a way that my judgment that it is a marsupial is not.

To illustrate this point, note that when I close my eyes, or when the kangaroo

steps momentarily out of my field of vision, it is no longer the case that I am seeing anything as a kangaroo. By contrast, if I form the judgment that a given animal is a marsupial (or indeed a kangaroo), this thought can persist in my awareness independently of my particular sensory state. Again, this marks a striking difference between any putative form of judgment involved in mid-level categorization and more canonical cases of non-perceptually integrated judgment.

Finally, note that the simple account described thus far may encounter problems when it comes to the apparent capability of many animals to learn new categories. As noted, these capacities do not seem readily explicable in terms of low-level categorization. However, if we seek to explain mid-level categorization in terms of full-blown propositional judgment, then prima facie, we must allow that animals capable of mid-level categorization are able to make such judgments about objects in their environment. While some philosophers are happy to grant this, it remains a minority view, especially in the case of simpler creatures such as bees.

The view given here – that mid-level categorization is a matter of rapid judgment – is of course quite simplistic, and there may be much more sophisticated ways one could spell out a cognitive account of mid-level categorization that elegantly captures the differences between mid-level categorization and cognitive processes like thought, judgment, and belief. However, it seems to me again that such accounts are liable to commit to a kind of pluralism about cognition that prompts the question of why it should be considered one core psychological capacity rather than many.

5. Mid-Level Categorization as a Third Kind

5.1. The Three Kinds view

The considerations given above are not intended as decisive arguments, and I remain entirely open to the possibility that our best theory of the mind may understand mid-level categorization as a case of perception or cognition proper. In this final section, however, I wish to explore a more radical alternative for explaining mid-level categorization that I term the Three Kinds view. In short, this proposal is that, in light of its prima facie differences from canonical cases of perception and cognition, mid-level categorization might be usefully considered a distinctive faculty in its own right.

Before spelling out the view, I should briefly clarify would it mean to consider mid-level categorization a faculty in its own right. Roughly, I take this to be a matter of explanatory utility. To the extent that cognitive science can avoid confusion and make better sense of the mind by separating mid-level categorization off from perception and cognition proper, then we will have grounds for treating it as a distinct psychological process.[9]

---

[9] One might wish to go even further than this, and suggest that perception, cognition, and mid-level categorization are in some sense psychological natural kinds; in the terms popularized by Ned Block (2014), a joint in nature between mid-level categorisation and the other psychological kinds.

With this in mind, I suggest that we should explore the view that mid-level categorization is a proprietary process that serves as an *intermediary* between perception and cognition, whose function is, roughly speaking, to classify the outputs of perception in terms of both innately specified and learned categories relevant to an organism's goals. Distinct from both perception and cognition, this faculty would thus take (perhaps nonconceptual) inputs from perception and assign them a conceptual or otherwise categorical attributive.[10]

This is the broad idea underpinning the Three Kinds view. Note that in claiming that mid-level categorization should be considered distinct from both perception and cognition, this formulation of the Three Kinds view constructively opens itself up to various challenges. For one, if it is found that distinguishing mid-level categorization from both perception and cognition glosses over philosophically or scientifically important differences, or bogs us down with hair-splitting minutiae of classification, it may seem like an unhelpful way of carving up the mind that fails to earn its explanatory keep.

5.2 – Evidence for the three kinds view

My purpose in what follows, then, is to press the claim that the Three Kinds account may indeed be able to earn its keep, insofar as it allows us to make sense of phenomena that sit uncomfortably within the traditional categories of perception and cognition, beginning with some broad theoretical considerations.

First, note that it seems well placed to avoid some of the complications raised in the previous section for narrowly perceptual or cognitive accounts. Thus, in the case of animal categorization, it may be the case that creatures which lack full-blown propositional thought nonetheless possess a faculty of mid-level categorization that enables them to assign simple conceptual or semantic labels to the outputs of perception. Similarly, on the natural supposition that the mechanisms underpinning mid-level categorization are fast and automatic, and that their outputs are at least somewhat resistant to correction by background beliefs, we can readily account for the existence of categorical forms of pareidolia discussed above.

The Three Kinds view is also well placed to accommodate a number of experimental results. For example, note that it provides a fairly straightforward account of associative agnosia, namely that the condition involves specific impairment of mid-level categorization, while leaving perception and cognition proper intact. Additionally, it seems readily able to account for the finding of Grill-Spector and Kanwisher discussed earlier that basic category information is present from the moment a subject becomes aware of a visual stimulus. On the reasonable assumption that subjects relied upon cognitive resources like working memory in actually *responding* to a given

---

[10] Note that the Three Kinds view can remain neutral on phenomenological issues, such as the rich content view (Siegel, 2010), and does not commit us to a position about what phenomenal character if any is evinced by the semantic representations processed in mid-level categorisation.

stimulus, we would naturally expect basic category information to be just as rapidly and readily accessible as low-level sensory information, at least insofar as we take mid-level categorization to operate prior to cognition proper. In contrast, high-level categorization would require a further cognitive step, namely some inference made on the basis of the outputs of low- and mid-level categorization.

A further important line of research that might bolster the Three Kinds view comes from emerging data in cognitive psychology concerning short-term memory mechanisms. Traditionally, debates about short-term memory have focused on nonconceptual sensory memory such as iconic memory and properly cognitive mechanisms such as working memory (Baddeley, 2003) and focal attention (Cowan, 2001). This has been reflected in philosophical treatments of perception and cognition, with a number of theorists (including Block, 2007, and Dretske, 1981) broadly identifying perception with sensory forms of memory and cognition with central cognitive processes such as working memory as noted earlier.

However, recent evidence suggests the existence of a third kind of short-term memory involved in our awareness of the world, namely conceptual short-term memory or CSTM. First proposed by Mary Potter, CSTM is "a mental buffer in which current stimuli and their associated concepts from long term memory… are represented briefly, allowing meaningful patterns or structures to be identified" (Potter, 2012).[11]

Note that current evidence suggests CSTM cannot readily be identified with working memory or sensory forms of memory such as iconic memory. In one crucial experiment, subjects saw 6-12 images presented in sequence for durations of 13, 27, 53, or 80ms (see Fig. 2). They were given a description of a target image (for example, 'wedding' or 'flowers') either 900ms or 200ms after seeing the images, and were asked to identify whether any images presented matched the description. On trials where the target stimulus was indeed present, subjects were given an immediate follow up recognition task to assess whether they could recognize the image when presented.

[11] Note that while Potter understands CSTM to be a properly *conceptual* mechanism, the results discussed below do not strictly require this. An alternate possibility is that information in CSTM in fact uses a nonconceptual but nonetheless categorical representational format. Additionally, Potter sometimes speaks of CSTM as a special form of working memory, whereas for clarity I have here distinguished it from working memory.

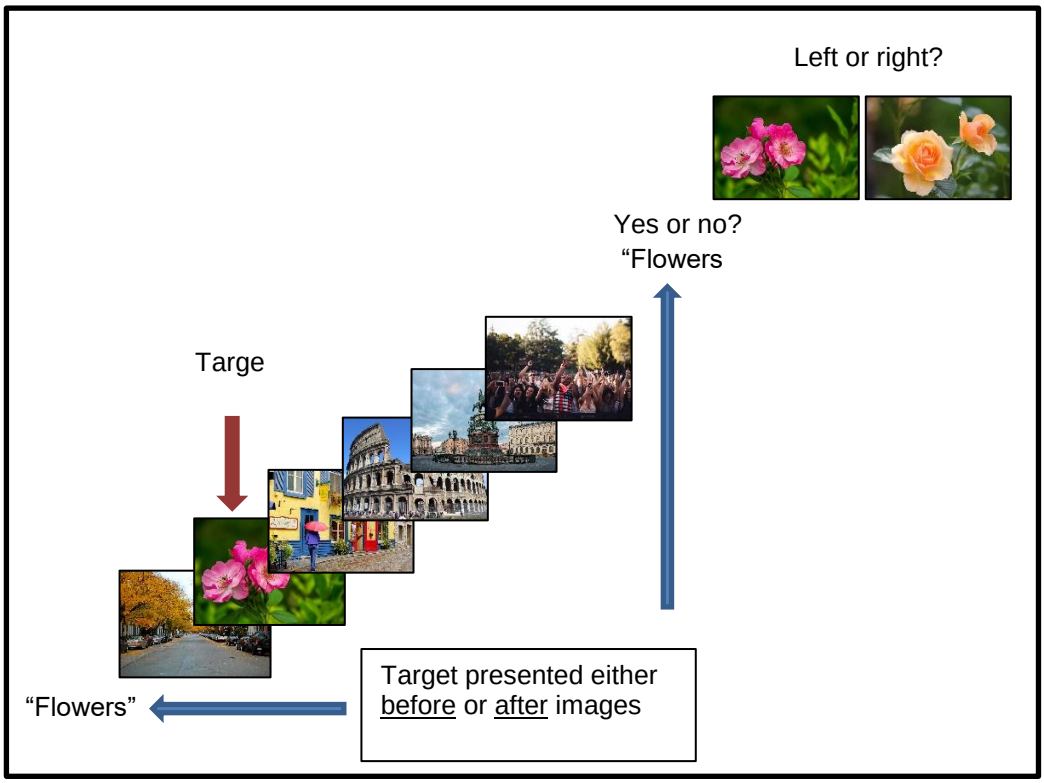

Fig. 3. Adapted from Potter et al. 2014.

Subjects displayed above-chance performance on all measures, and in the follow up recognition task they were consistently able to recognize the target only if they had already detected its presence under the cued description. Crucially for present purposes, the results suggest that subjects were able to briefly retain categorical information about all or most of the 6-12 presented images. Otherwise, it is hard to see how they would be able to accurately report on the presence or absence of a given target when cued immediately *after* their presentation. This suggests that they were not merely storing the relevant information in working memory, given that this is generally taken to be a strictly capacity limited mechanism (see, e.g., Cowan 2001).

However, it is also unlikely that this information was stored in sensory memory: such buffers are normally taken to encode strictly sensory properties of stimuli such as color, size, and shape, and to be overwritten by sequential presentation of stimuli, whereas the tasks in this experiment seemingly required subjects to retain fairly high-level categorical information about images presented one after another.[12] These considerations lead Potter to claim that the results provide evidence for an intermediate high-capacity buffer, CSTM, in which the categorical identities of presented images are (very briefly) encoded.

For present purposes, note that CSTM seems to exhibit many of the functional properties of mid-level categorization as described above. Specifically, it is fast, automatic, has a brief duration, and serves to encode sensory information in respect of

---

[12] There are of course several alternate debunking explanations for the results. For a more detailed discussion of these, see Shevlin (2017).

learned categories. It may thus constitute a promising locus for mid-level categorization within the mind's different systems, and one that is plausibly distinct from the processes taken by many philosophers to underlie perception and cognition in the strict sense.

Of course, even if we accept that mid-level categorization involves a distinctive kind of short-term memory, it does not follow that it cannot be a special case of perception or cognition. For example, perception might be realized across multiple sequential forms of memory. Nonetheless, to the extent that these different memory stores have different functional properties such as capacity, duration, and representational format, our cognitive science is likely to benefit from distinguishing between the psychological processes that exploit them, thus bolstering the case that mid-level categorization is worth regarding as a third kind in its own right.

A further experiment involving rapid-short term categorization that also arguably lends weight to the Three Kinds view is the finding of Belke et al. (2008) that subjects are seemingly able to categorize more items at once than can be encoded in working memory. Subjects were presented with arrays of varying sizes (4 or 8 items) having been cued before presentation to look for a category target like 'motorbike' or 'cigar'. On certain trials this target was absent, but a semantically-related distractor (such as 'motorbike helmet' or 'pipe') was present. Using an eye-tracking camera, Belke et al. found that subjects' gaze was immediately drawn to both targets and distractors, and that this occurred regardless of array size. Additionally, they found that subjects' eye movements were just as reliably drawn to the targets and distractors even under a cognitive load condition (though their responses were delayed).

Insofar as subjects' gaze was reliably attracted to semantically-relevant items under larger 8-item arrays and under conditions of cognitive load, this experiment provides support for the idea that the initial extraction of relevant categorical information does not to rely just on capacity-limited working memory. In this sense, it is an immediate complication for narrowly cognitive views of mid-level categorization, at least insofar as these views wish to associate cognition with working memory. By contrast, the finding is quite compatible with both the perceptual account of mid-level categorization and the Three Kinds view under discussion. However, note that to the extent that perceptual theorists are inclined to identify perception with mechanisms such as iconic memory, this result may present a challenge as there is reason to think that iconic memory does not encode categorical information (Sperling, 1960), and Belke et al. specifically invoke CSTM to explain subjects' capacities.

5.3 – The representational format and the Three Kinds view

The above arguments are not intended to show that the Three Kinds approach is straightforwardly *preferable* to an account that explains mid-level categorization just in terms of perception or cognition proper. However, I hope I have provided some initial reasons for thinking that the view might be worth exploring.

Before closing, I wish finally to consider one complication for the Three Kinds view. This concerns the content and format of representations involved in mid-level categorization. As noted earlier, there is reason to think that low- and mid-level

categorization have different kinds of content (roughly, analogue and digital) and are subserved by vehicles with a different representational format. However, I have not yet made any commitments concerning the kinds of content and format we might expect in mid-level categorization Given that a number of theorists (including Dretske, Block, and Burge) explicitly appeal to representational format to distinguish perception and cognition, this may seem a worrying oversight. In particular, some theorists might suggest that, if categorical perception involves representations with a propositional structure, it is *ipso fact*o part of cognition; by contrast, if it involves iconic representations derived from perceptual inputs, it is straightforwardly part of perception.

I would resist this move on two grounds. First, note that representational format is not the only basis on which to distinguish psychological processes: considerations such as function, phenomenology, and mechanism are also relevant. Thus, I see no difficulty, for example, in someone's adopting the Three Kinds view while simultaneously holding that the representations involved in categorical perception had propositional structure and content.

However, there may also be grounds for thinking that considerations about representation format can be used to further distinguish mid-level categorization from both cognition and perception proper. In particular, I am drawn to a view according to which the outputs of the mechanisms of mid-level categorization have a *hybrid* format, in which a conceptual or otherwise semantic label is indexed to an underlying iconic representation.

Such hybrid representational formats are familiar to us in the form of labelled maps and diagrams (Camp, 2007). As a simple example, consider the seating chart shown below. Here, we can exploit the symbolic elements of the representation to conclude that there are more women than men around the table, but we can also use the analogue spatial structure of the chart to derive a large number of more specific conclusions, for example that a man in the top-left hand corner is seated diagonally opposite another man, that the man to his right is sitting further back, and so on.

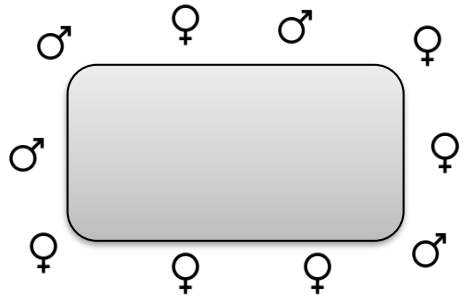

Fig. 4. A simple example of a hybrid representation showing men and women seated at various point around a table.

While I claim that the format of categorical perception involves a semantic element, I wish to leave open whether this element possess properly *conceptual* content. I am inclined to regard this as a matter that depends on one's broader theoretical commitments about concepts. At a minimum, however, I would suggest this semantic

component is a repeatable, attributive with a wide representational range encompassing learned categories. Additionally, unlike iconic representations, it need not rely whatsoever on structural isomorphisms with its representata in order to function representationally.

This may fail to satisfy more demanding accounts of concepts. The semantic component in categorical perceptions may not, for example, have content that is freely recombinable in a way that satisfies Evans' generality constraint (Evans, 1982). Additionally, I leave open the possibility that it may not be embedded within a full-blown propositional structure, but instead be bound to an underlying sensory representation, with the result that it would fail to count as a concept by the lights of Burge (2010). I am happy to leave such questions open for present purposes.

The appeal to a hybrid format allows us to more clearly distinguish mid-level categorization from what I take to be the plausibly iconic format of perception and the full-blown propositional structure of judgments and beliefs. Moreover, this view is arguably also independently motivated, insofar as it captures the phenomenological aspects of mid-level categorization and the way in which the kind of awareness accompanying mid-level categorization is integrated with awareness of lower level perceptual properties (as described in 4.1, above). Thus, in viewing the pareidolic images given earlier, I am aware of the items not merely as an eye and an owl, but an eye and an owl with specific color, shape, and size.

Similarly, a hybrid account seems well placed to accommodate the experimental data of Potter et al. concerning CSTM. Recall that subjects were only able to distinguish a presented image from another with the same rough semantic content if they had first detected it under the relevant description. In other words, it seems that the process of categorizing the image appropriately facilitated a subsequent task (namely, recognition) in which the sensory features of the image were relevant. This is precisely what one might expect were categorical perception to involve representations with a hybrid format, in which encoding the semantic identity of a target also involved encoding at least some of its sensory features.

Hence I suggest that the Three Kinds view has much to recommend it. While, as noted, I do not discount purely perceptual or cognitive accounts of mid-level categorization, the idea that mid-level categorization is a faculty in its own right is one worthy of further elaboration and investigation.

(6) Conclusion

My primary goals in this paper have been threefold. First, I sought to suggest a three-part framework for thinking about the varieties of categorization at work in our awareness of the world. Second, I suggested some ways in which we could understand mid-level categorization as a special case of perception or cognition. Finally, I outlined an alternative proposal, the Three Kinds view, that takes mid-level categorization to be a process independent of both perception and cognition. This final proposal is tentative, of course, but I believe it may offer a promising and hitherto unexplored framework for

understanding the place of mid-level categorization within the mind.

REFERENCES

Baddeley, A.D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829-839.

Bayne, T., & Montague, M. (2011). Introduction to Cognitive Phenomenology. In Bayne, T., & Montague, M. (eds), *Cognitive Phenomenology*. Oxford University Press UK.

Beck, J. (2015). Analogue Magnitude Representations: A Philosophical Introduction. *The British Journal for the Philosophy of Science*, Volume 66, Issue 4.

Belke, E., Humphreys, G., Watson, D., Meyer, A. and Telling, A., (2008). Top-down effects of semantic knowledge in visual search are modulated by cognitive but not perceptual load. *Perception & Psychophysics*, 70 8, 1444 – 1458.

Block, N. (2007). Consciousness, Accessibility, and the Mesh Between Psychology and Neuroscience, *Behavioral and Brain Sciences* 30, pp. 481–499.

Block, N. (2014). Seeing-As in the Light of Vision Science. *Philosophy and Phenomenological Research, 89*(3), 560-572.

Block, N. (2016) Tweaking the Concepts of Perception and Cognition. *Behavioral and Brain Sciences* 39, 21-2.

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

Brooks, D. I., Freeman, J. H., & Wasserman, E. A. (2013). Categorization of photographic images by rats using shape-based image dimensions. *Journal of Experimental Psychology: Animal Behavior Processes*, Vol 39(1), Jan 2013, 85-92.

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

Byrne, A. (2009). Experience and content. *Philosophical Quarterly* 59 (236):429-451.

Camp, E. (2007). Thinking with maps. *Philosophical Perspectives* 21 (1):145–182.

Carey, S. (2009). *The Origin of Concepts*. Oxford University Press.

Carruthers, P. (2009). Invertebrate concepts confront the generality constraint (and win). In R.W. Lurz (ed.), *The Philosophy of Animal Minds*. Cambridge University Press 89-107.

Carruthers, P. (2014). On central cognition. Philosophical Studies 170 (1):143-162.

Chittka, L., & Jensen, K. (2011). Animal Cognition: Concepts from Apes to Bees. *Current Biology, 21*(3)

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

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences 24*1, 87-114.

Dretske, F. (1981). *Knowledge and the Flow of Information*. MIT Press.

Dretske, F. (2004). Change Blindness. *Philosophical Studies* 120, 1-18.

Duffy, C.J. (1999). Visual loss in Alzheimer's disease: out of sight, out of mind. *Neurology*. 52 (1): 10–1.

Evans, G., 1982. *The Varieties of Reference*, Oxford: Oxford University Press.

Fodor, J. (1983). *Modularity of Mind*. MIT Press.

Fodor, J. (2007). The revenge of the given. In Brian P. McLaughlin & Jonathan D. Cohen (eds.), *Contemporary Debates in Philosophy of Mind*. Blackwell 105--116.

Fodor, J. (2015). Burge on Perception, in E. Margolis and S. Laurence (eds.), The Conceptual Mind: New Directions in the Study of Concepts. Cambridge, MA:MIT Press.

Grill-Spector, K., & Kanwisher, N. (2005). Visual Recognition. *Psychological Science*, 16(2), 152-160.

Kirkpatrick, K., Bilton, T., Hansen, B. C., & Loschky, L. C. (2014). Scene gist categorization by pigeons. *Journal of Experimental Psychology: Animal Learning and Cognition, 40*(2), 162-177.

Lupyan, G. (2015). Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of Philosophy & Psychology* 6 (4):547-569.

McDowell, J. (1994). *Mind and World*. Cambridge: Harvard University Press.

Peacocke, C. (1992). *A Study of Concepts*, Cambridge MA: MIT Press.

Pitcher, G. (1971). *Perception*, Princeton: Princeton University Press.

Potter, M.C., Wyble, B., Hagmann, C.E., & McCourt, E.S. (2014): Detecting meaning in RSVP at 13ms per picture. *Attention, Perception, & Psychophysics*, 762, 270-279.

Potter, M. C. (2012). Conceptual Short Term Memory in Perception and Thought. *Frontiers in Psychology*, 3:113.

Prinz, J. (2011). The Sensory Basis of Cognitive Phenomenology. In T. Bayne, & M. Montague, (eds.), *Cognitive Phenomenology*. OUP.

Quilty-Dunn, J. (2016). Iconicity and the Format of Perception. *Journal of Consciousness Studies* Vol. 23 (3-4):255-263.

Rescorla, M. (2009). Chrysippus' dog as a case study in non-linguistic cognition. In Robert W. Lurz (ed.), *The Philosophy of Animal Minds*. Cambridge University Press. pp. 52--71.

Shevlin, H. (2017). Conceptual Short-Term Memory: A Missing Part of the Mind? *Journal of Consciousness Studies*, Vol. 24, No. 7-8, July/August 2017.

Siegel, S. (2010). *The Contents of Visual Experience*. Oxford.

Sperling, G. (1960). The Information Available in Brief Visual Presentations, *Psychological Monographs: General and Applied* 74, pp. 1–29.

Stazicker, J. (2018). The Visual Presence of Determinable Properties. In *Phenomenal Presence*, eds. Dorsch, F., and Macpherson, F. OUP.

Tye, M. (1995). *Ten Problems of Consciousness*. Cambridge, MA: MIT Press