

## **Distinguishing Top-Down From Bottom-Up Effects**

**Nicholas Shea**

### **Abstract**

The distinction between top-down and bottom-up effects is widely relied on in experimental psychology. However, there is an important problem with the way it is normally defined. Top-down effects are effects of previously-stored information on processing the current input. But on the face of it that includes the information that is implicit in the operation of any psychological process – in its dispositions to transition from some types of representational state to others. This paper suggests a way to distinguish information stored in that way from the kind of influence of prior information that psychologists are concerned to classify as a top-down effect.

So-drawn, the distinction is not just of service to theoretical psychology. Asking about the extent of top-down processing is one way to pose some of the questions at issue in philosophical debates about cognitive penetrability – about the extent of the influence of cognitive states on perception. The existence of a theoretically-useful perception-cognition distinction has come under pressure, but even if it has to be abandoned, some of the concerns addressed in the cognitive penetrability literature can be recaptured by asking about the extent of top-down influences on any given psychological process. That formulation is more general, since it can be applied to any psychological process, not just those that are paradigmatically perceptual.

### **1 Introduction**

The distinction between top-down and bottom-up effects plays a central role in experimental psychology. But there is a problem with the way it is standardly drawn. Even if there were a completely Fodorian perceptual module, fully encapsulated against the effects of any beliefs, desires, or other representational states found elsewhere in the psychological architecture, it would still rely on its own pre-existing store of information. After all, the idea of a Fodorian module, as opposed to a mere reflex, is of a genuine information processing system, one in which the transitions between representations within the module embody or encode information that is appropriate to the circumstances in which it operates. A visual module that uses a contrast map to calculate the edges of objects embodies the information that contrast boundaries are a reasonably reliable sign of edges. These inbuilt expectations, implicit in the

operation of the module itself, are rightly not considered to be top-down effects. Yet they exemplify the impact of pre-existing information in the system on the operation of a perceptual process. If the question of the relative influence of top-down information is to be well-posed, we need to show that the top-down influence of antecedent information can be distinguished from the proprietary information embodied in a module.

The purpose of this paper is to say something constructive – although by no means conclusive – about how top-down effects on the operation of a psychological process should be distinguished from the influence of the information implicit in the operation of the process. So-formulated, the distinction between top-down and bottom-up effects can be put to philosophical work. It allows us to ask, for any psychological mechanism, how much its operation is driven by current input and to what extent it is influenced by prior information. Answering that question will form an important part of spelling out how sensory information is processed. It also provides a basis for asking epistemological questions about the relations between sensory systems and the world, relations amongst sensory systems, and their relations to other psychological capacities. These issues also arise in the philosophical literature on the cognitive penetrability of perception by cognition. The top-down vs. bottom-up distinction allows us to pose some of those questions more generally, not just about paradigmatically perceptual states. So the distinction would allow some of the concerns in the cognitive penetrability debate to be preserved, even if a distinctive category of the perceptual turned out to be theoretically unsustainable.

Section 2 introduces the top-down / bottom-up distinction and sets up the problem. Section 3 uses the distinction between explicit and implicit representations to address the problem. Section 4 shows that, even if it is problematic to think that there is a distinctive category of the perceptual, the top-down / bottom up distinction is suited to characterising some philosophical issues about the balance between current input and pre-existing information that arise in the cognitive penetrability literature.

## **2 Top-down vs. bottom-up**

A central concern about the senses is to understand the extent to which a process that uses current input to track current and potentially changing features of the immediate environment is also influenced by antecedently-represented information. In scientific psychology that question is approached by asking to what extent a psychological process is driven by top-down as opposed to bottom-up information. Here is how it is presented in standard psychology texts:

‘Bottom-up processing is processing which depends directly on external stimuli, whereas top-down processing is processing which is influenced by expectations, stored knowledge, context and so on.’ (Eysenck, 1998, p. 152)

‘In bottom-up processing (also called data-driven or stimulus-driven processing), the process starts with the features – the bits and pieces – of the stimulus, beginning with the image that falls on the retina. This information is processed hierarchically by successively higher levels of the visual system until the highest levels (the “top” of the system) are reached, and the object is perceived. Top-down processing (also called

knowledge-driven processing) involves the use of contextual information supplied from memory – the “big picture”.’ (Carlson et al., 2010, p. 202)

The first quotation brings out the essential difference between information directly drawn from a stimulus and information drawn from memory. The second quote illuminates why the metaphors of “bottom” and “top” are used – they come from an assumed hierarchy of information processing. To the extent that antecedently stored information has an effect on processing, it is assumed to derive from higher levels in this hierarchy. That gloss is subsidiary. The core is the distinction between incoming and pre-existing information.

Two other connotations of the term should be set aside. First, especially in cognitive neuroscience, top-down is sometimes understood in neural terms. Top-down effects are those that proceed causally from ‘higher’ to ‘lower’ neural areas (Mechelli et al., 2003), where there is a pre-existing conception of which brain areas are higher and which lower in the hierarchy (very approximately: moving forward from the back of the brain, with the prefrontal cortex at the top of the hierarchy, but with the motor cortex somewhat anomalous in that it lies behind the prefrontal cortex but comes later in information processing and is more closely tied to the periphery than prefrontal, and most temporal and parietal areas).

The distinction is not at base a neural one. The difference between higher and lower neural areas itself derives from a rough model of where various areas fall in an information processing hierarchy. Visual information proceeds from the retina, through the LGN and is first processed cortically in the occipital lobe. Visual information flows through a hierarchy of visual areas in the occipital cortex before spreading to the temporal lobe and the parietal lobe (usually thought to occur in parallel, in ventral and dorsal visual streams). It may then be used directly to drive the motor cortex and drive action (in a sensorimotor loop), but that process may be controlled or modulated by information processing in the prefrontal cortex at the top of the hierarchy. There are lots of places to object to this parody of the hierarchy of information processing in the brain, but something like this underlies the idea, to take one example, that connections from prefrontal or parietal cortex to the occipital lobe carry top-down information. The categorisation of a neural area as higher or lower derives from the information processing hierarchy in which it is located. So the neural distinctions are not constitutive of, but merely evidence of, the difference between top-down and bottom-up effects.

A second connotation is that top-down influences are voluntary, or derive from the self, the will or the all-purpose homunculus that is sometimes hidden behind the label ‘executive control’. That gloss is not important for our purposes here, so I will set it aside – top-down effects need not emanate from some special locus of endogenous or voluntary control, if indeed there is such an entity, which is doubtful.

The psychologists’ distinction is drawn in terms of causal effects of prior representations, but not all causal effects are relevant: only those that occur within information processing. A prior belief might causally influence perceptual processing by causing a subject to move her eyes, thus changing the input. Then visual processing would have been ‘influenced by stored knowledge’, but not in the way psychologists were thinking of. Their concern is with cases where prior representations have a representational influence on the processing of input –

with influences that occur computationally, that is to say, either directly, or via a chain of other representations.<sup>1</sup> Causal influences mediated by the world do not count.<sup>2</sup>

The psychologists' distinction between top-down and bottom-up effects can be used to frame two philosophical questions about any psychological process. Firstly, in characterising a psychological process (for example here, a sensory process), we can ask how much the process is driven by incoming information from external stimuli, and how much it is affected by antecedent representations that were already in place before the stimulus was encountered. Secondly, there is an epistemological question. We can ask how the outcome of any piece of processing M1 is suitable for belief formation or any other piece of subsequent processing M2, if antecedent representations in M2 have affected the output of M1. It is not just the effect of cognition on perception that raises the risk that we may be illegitimately pulling ourselves up by our own epistemic bootstraps. Any kind of circular relations between earlier and later processing raises questions about the reliability of the process (Lyons, forthcoming). An appropriate account of the epistemic profile of a sensory process – its sensitivity, specificity, and positive and negative predictive value – will depend upon the extent to which there are top-down as well as bottom-up influences on that process.

That is not to say that top-down influences are necessarily epistemically pernicious. On Karl Friston's comprehensive model of the brain, feed forward signals consist only of prediction errors (Friston, 2010; Friston & Stephan, 2007). It is the top-down signals that represent what is the case. In this way top-down signals directly affect, or even constitute, what you represent, but without creating a self-reinforcing cycle, because it is not the result of these predictions that is fed forward as input to subsequent processing, but only the difference between prediction and current input.<sup>3</sup> Similarly, in the closely-related Bayesian models of neural processing, priors (antecedent representations) strongly constrain the way incoming information is processed (and hence what is perceived), but because of Bayesian updating, the priors don't just end up confirming themselves in a cascade of illusory justification – they are gradually updated in the direction of the incoming evidence. Of course, false priors can compromise veridical perception, and if you start with very bad priors you can end up misperceiving for a very long time. But that is not enough to show that the influence of prior representations on what you perceive violates an epistemic norm. Indeed, Bayesians think that reaching hypotheses about what is the case by combining antecedent beliefs with incoming information via a Bayesian updating rule constitutes compliance with an epistemic norm.

It is not only on Friston's model that top-down influences are seen to be epistemically acceptable. In general, the question of how top-down influences affect the reliability of a psychological process should be considered case-by-case, in the light of information about what the top-down effects are and how they operate. Some top-down influences, including from background beliefs, may be helpful in narrowing the range of expected stimuli or changing the priors about what is likely to be encountered, in a way that increases detectability,

---

<sup>1</sup> That is to understand computation in a broad sense, not restricted to classical computation.

<sup>2</sup> Thanks to James Stazicker for pressing me to make this explicit.

<sup>3</sup> Structurally the same point is true of the interplay between feedforward and feedback processing in Eliasmith and Anderson's control-theoretical model of the brain (Eliasmith & Anderson, 2003).

discriminability or processing speed. Others may reduce the reliability with which features of the world are represented, especially in novel contexts. The impact on reliability will also depend on the kinds of environments under consideration, in which the psychological process operates. In short, although the influence of top-down effects on the online processing of current input need not in principle be epistemically pernicious, whether it is remains an important question that can only be properly answered in the light of knowing how bottom-up and top-down information combine to drive the information processing involved. Furthermore, a normative account of the epistemic relations between various representational states should be informed by the way that top-down and bottom-up information are in fact balanced by the actual psychological systems in question.

So the psychologists' category of top-down effects is well-suited to formulating some central philosophical concerns about sensory and other psychological systems. There is a problem, however, with the psychological distinction. Top-down effects are defined as influences of an individual's expectations and stored knowledge (Eysenck, 1998, p. 152). On the face of it, that includes the expectations that are implicit in the operation of a psychological process – in its dispositions to transition from one representation to another. Perceptual learning, for instance, affects the way processing within a perceptual module is disposed to unfold (Fahle & Poggio, 2002). As a result of perceptual learning the system will process input in a different way. The new transitions it is disposed to make will embody a different set of expectations. But the process does not do so by drawing on a separate source of information antecedently represented elsewhere. Indeed, the information implicit in the operation of a perceptual process, resulting from perceptual learning, is not explicitly represented anywhere.

For example, the ability to discriminate visual gratings with different luminance distributions improves with experience, but the effect does not transfer between vertical and horizontal gratings, suggesting that changes relatively early in visual processing are responsible for the improvement (Fiorentini & Berardi, 1980). Similar effects are found in many other domains such as visual acuity (Fahle, Edelman, & Poggio, 1995), and discriminating the direction of visual movement (Watanabe, Nanez, & Sasaki, 2001). Neuroimaging results suggest that the plasticity responsible for some of these changes occurs within the neural areas that carry out the early stages of processing (Maertens & Pollmann, 2005). Neurophysiology confirms that some of the changes in processing categorised as visual adaptation, on shorter timescales up to a few minutes, are due to changes within, rather than influences on, the processes constituting early sensory processing (Kohn, 2007). Information processing models show how such effects can be achieved by changes to the way a psychological process is disposed to transition between various representational states (Poggio, Fahle, & Edelman, 1992).

These kinds of changes in the online processing dispositions of a psychological process do not only occur within low level perceptual processes. Another large class of results concern sensorimotor learning. Rather than asking subjects to make a perceptual discrimination which is reported in some unnatural manner, like pressing a button, sensorimotor experiments are more dynamic, requiring subjects to mediate between stimulus and action in a fluid way (like visually-guided reaching for a potentially moving target). These sensorimotor loops also undergo rapid adaptation or tuning in the light of experience (Mazzoni & Krakauer, 2006), often mediated by primary sensory or motor cortices, or the cerebellum (D'Angelo & De Zeeuw, 2009). A sensorimotor loop, consisting of a set of dispositions to connect perceptual

input with motor behaviour, can be tuned by experience by modifying those dispositions via synaptic plasticity, without that being mediated by the effect of some separate explicit representation on the system (De Zeeuw et al., 2011).

By contrast, in other cases there is evidence that relatively early cortical systems processing sensory input are affected by information antecedently represented beyond those systems (Di Lollo, Enns, & Rensink, 2000; Summerfield & Koechlin, 2008; Ulzen et al., 2008) – although such claims remain controversial. Of course, it is by no means settled that all, or even any, of the effects described in the large literatures on perceptual learning and sensorimotor adaptation are definitely the result of changes to the dispositions embodied in a psychological process, rather than being mediated by top-down influences from other parts of the system. But the important thing for our purposes is that these debates are coherent. The question asked in the perceptual learning literature is whether a piece of perceptual learning consists merely in a modification in the dispositions to respond bottom-up to stimuli, or whether instead it depends upon top-down influences. For that question to be well-posed the contrast must reflect a genuine distinction. The problem is that such cases fall within the letter of the psychological definition of top-down effects, while falling outside the spirit that the distinction is aiming to capture. That would spell trouble for placing reliance on the top-down / bottom-up distinction to formulate philosophical questions, unless the distinction can be clarified to exclude these cases in a principled way.

### 3 Explicit vs. implicit

To recap: to make it legitimate to appeal to the top-down / bottom-up distinction for our philosophical purposes, we need to show that the top-down influence of antecedent representations can be distinguished from the information that is implicit in the set of inferential dispositions embodied in a psychological process. There is an obvious way of drawing that distinction, which doesn't quite work, but can be turned into a tenable solution.

The obvious distinction is between the occurrent and dispositional senses of representation. When beliefs, desires or other mental states affect the processing of sensory information, they are occurrent, and their being tokened is necessary for there to be an effect on sensory processing. By contrast, the information embodied within a perceptual module or other piece of psychological processing exists in virtue of dispositions – the disposition of the process to move from certain representational states to others (e.g. from such-and-such arrangement of visual contrast to representations of edges). Can't we just say that top-down effects occur only where *occurrently represented* pre-existing expectations, beliefs, memories, etc. influence the way incoming information is processed?

Not quite, because of a difficulty with construing the occurrent / dispositional contrast. A representation becomes occurrent when it is tokened – when there is a physical change such that an instance of that representation type is realised in such a way that it can have a causal influence on subsequent psychological processing. The difficulty arises because there is, of course, a causal basis for perceptual learning, sensorimotor adaptation, and so on. In general, the dispositions that a psychological process has, to transition between certain types of representational states, will have a physical basis. When those dispositions change, it is because the physical basis changes. Since it is legitimate to think of those dispositions as embodying expectations or containing information, the causal basis of the disposition is the causal basis of

the storage of that information. That furnishes a sense in which this apparently dispositional information is actually being tokened; which is not yet to show that there is a representation with that information as its content. But it does show that we need to do more than simply rely on the difference between tokened information and mere dispositions.

Fortunately, there is a relatively straightforward way to side-step this difficulty. We can make use of one final distinction: between implicit and explicit representations. These labels are used to mark many different contrasts (e.g. conscious vs. unconscious, verbally-reportable vs. not), but there is a rather tightly-delineable sense that will serve our purposes. The expectations or information embodied in the dispositions of a psychological process to transition from one representation to another are implicit, in the sense that they can have no impact on subsequent processing except via the representations which they connect. An explicit representation can potentially have an impact on many pieces of downstream processing. Consider, for example, a feedforward system that was disposed to make transitions from a light contrast map to representations of where the edges of objects lie. Those inferential dispositions implicitly encode, *inter alia*, the information that the edges of objects tend to occur at the spatial locations where discontinuities in contrast levels are found. That information has an effect on subsequent processing, but that influence is wholly in virtue of its effect in producing the representations of edges. By contrast, the representation of the location of edges is an explicit occurrent representation. It can act as input to many different systems: object discrimination, object categorisations, online guidance of reaching, and so on.

One clear illustration of this contrast is in the way information is stored and acted on by a connectionist network. A connectionist network has a set of dispositions to transition from patterns of distributed activation at its input layer, through patterns of activation at hidden layers, to a pattern of activation at its output layer. The complete set of its dispositions to make transitions between these occurrent representations is fixed by the network's weight matrix. As a result of training, the network can acquire dispositions to transition from input to output that encode information. For example, the network might have a disposition to respond to representations of letters at the input layer by transitioning to a phonetic space at the hidden layer (including a contrast between vowels and consonants), and then onwards to a classification of the input by phonetic features at the output layer. Those inferential dispositions embody information about relevant statistical patterns in the material on which the network was trained. In that sense, the weight matrix itself encodes information. But that information is not explicit in the weight matrix or in the set of dispositions to move between distributed patterns of activation. The weight matrix only has an effect on downstream processing via the occurrent representations, dispositions to transition between which it underpins. The pattern of activation at the output layer, or a hidden layer, could act as input to many different subsequent pieces of processing. The weight matrix itself is not available to guide processing in that way. The information contained in the weight matrix is effective only in the way that it underpins dispositions to transition between explicit representations, from input through hidden to output layer.

So there is no need to deny that there is a sense in which the information embodied in the disposition to transition between representations in a particular way is itself realized – or tokened – in the causal basis of those dispositions. Rather than trying to argue that these are not occurrent for some reason, we can simply rely on the fact that this information is not

explicit. Only explicit representations can act as input to further computations. The information which is implicit in the dispositions embodied in a psychological process cannot act as input to further computations. Its only impact on information processing is through the dispositions which it underpins to transition between occurrent representations.

When an early perceptual process learns to discriminate the signs of a luminance boundary more finely, or when a sensorimotor process re-learns a mapping between visual input and motor output, the dispositions to transition between representations at various stages of processing are altered, but without the new information embodied in the process being explicitly represented anywhere, either to drive learning, or as a result of learning. (At least, that is one of the positions that is being argued for in debates about perceptual learning.) So asking whether the influence on psychological processing is merely implicit, or whether it is caused by some explicit representation outside that process, is a good way of distinguishing modifications to a psychological process from top-down effects on the process. The implicit-explicit distinction we have drawn allows us to pick out the category of top-down effects as intended by psychology, distinguishing it from perceptual learning, sensorimotor adaptation, and the like.

To summarise, I would argue for using terms as follows:-

*Occurrent representation*

The tokening of a representation, i.e. its realization in such a way that it can have a causal impact on psychological processing.

*Implicit representation*

A disposition to transition between two or more occurrent representations that can have no influence on subsequent processing except via the representations between which the disposition subsists.

*Explicit representation*

Occurrent representation that is not implicit.

*Top-down influence*

A representationally-mediated effect of an explicit representation R on a psychological process, where R is not computed more directly than the representational influence of current sensory input on the process.

*Bottom-up influence*

An effect on the outcome of a psychological process which is not a top-down influence.

This way of clarifying the use of terms does not clear up all difficulties, of course. In particular, I use a distinction between more and less direct computational routes, which could profitably be clarified further. But progress has been made: we have replaced the problematic distinction between the perceptual and the cognitive with a collection of better-understood psychological properties. Once we set aside representations that are merely implicit in dispositions to transition between various occurrent representations, the cash value of the



distinction between top-down and bottom-up influences is a matter of directness of computational influence. For any psychological process whose output is relatively directly influenced by current sensory input we can ask whether the output is also affected by more indirect routes: either by antecedent representations whose tokening is not caused by the current stimulus, or by representations whose connection to current sensory input is more indirect than the influence of sensory input on the process itself. Relative extent of top-down and bottom-up influence will certainly be a matter of degree. Furthermore, whether an effect counts as top-down at all may also be a matter of degree, if the relative directness of the influence of sensory input on a psychological process is a matter of degree. But graded distinctions are tractable here, provided it is reasonably clear what the gradations depend upon.

The top-down / bottom-up distinction has other merits. It admits of top-down effects of one sensory process on another, as well as top-down effects of beliefs and desires on sensory processing. Epistemic questions arise whenever the processing of sensory input is constrained by pre-existing representations, even when those representations derive from processing of the same stimulus some milliseconds earlier. A related merit is that the distinction does not require a strict hierarchy of information processing (*pace* the metaphorical use of “top” and “bottom”). The relations between psychological processes may be overlapping, parallel and intertwined in complex ways, but we can still in principle ask about the relative directness of two routes of influence on a given process. Furthermore, the distinction fits with the existence of “vertical” sensorimotor loops that directly mediate between sensory input and motor output in a dynamic, continually-adjusting way. The influence of beliefs, desires or other antecedent representations in modulating the targets for or other properties of these sensorimotor loops fits within our framework – it would be counted as a top down effect. In short, the top-down / bottom-up distinction can be drawn in an empirically-tractable way that bypasses the objection we raised at the outset.

#### 4 The Perception-Cognition Distinction

We have argued so far that the top-down / bottom-up distinction, when properly drawn, is a useful way to formulate important questions about the relative influence of incoming and pre-existing information on the operation of any psychological process. That is one of the issues at the heart of the philosophical literature on ‘cognitive penetrability’ – where it is formulated as the question of whether cognitive contents directly affect perceptual processes. Siegel defines the cognitive penetrability of visual perception as the nomological possibility that cognitive or affective states can cause a change in the visual contents that are or would be experienced while seeing and attending to the same distal stimuli under the same external conditions (Siegel, 2011, pp. 5-6).<sup>4</sup>

Two issues from the cognitive penetrability literature correspond to the questions we discussed above. First, it is thought to be central to an adequate characterisation of perceptual experience to identify the extent to which it is, or can be, penetrated by the contents of belief.

---

<sup>4</sup> For the case of desire rather than belief, Stokes’ (2011) *orectic penetration hypothesis* is the claim that desires or other desire-like cognitive states causally influence perceptual states via an internal mechanism.

That is the correlate of our more general question as to how any psychological process is influenced by the pre-existing representations explicitly represented in any other psychological process. Within the literature on perception this answers the question: how much is perceiving just a matter of receptivity to the outside world, and how much is it a constructive process based on what we antecedently represent? Secondly, the cognitive penetrability literature is interested in the epistemological question mentioned above. If the operation of our senses is systematically affected by what we want or already believe, the epistemological project becomes more challenging. How then can perceptual contents justify beliefs? There is reasonable evidence that such cases occur, yet they make trouble for various accounts of the normative relations between perceptual contents and doxastic states (Siegel 2011). Getting clear about these two issues will be important for a full understanding of the senses.

Siegel's definition reflects a widely-shared assumption in the literature on cognitive penetrability or cognitive penetration, in that it presupposes that there is a theoretically-important distinction to be drawn between perception and cognition (Macpherson, 2011; Stokes, 2011). That assumption may be justified given the central role played by the idea of the perceptual in folk psychology. But it has come under pressure, so it would be useful if some of the concerns of those interested in cognitive penetrability could be raised without presupposing that the perceptual forms a distinctive category that differs in theoretically-important ways from other forms of psychological processing.

This section sets out some problem cases which suggest that it may be more difficult to individuate the category of the perceptual than has been supposed. The point is not just that it is currently unclear where to draw the line; nor just that there are borderline cases; nor that a representation's being perceptual is a matter of degree. Rather the worry is that the important issues here, especially concerning the relative balance between input and pre-existing representations and the epistemological consequences thereof, arise in just the same way for a range of broadly input-driven systems that are not paradigmatically perceptual. That suggests that there may be little merit in identifying a supposedly special class of perceptual representations and investigating their supposedly proprietary epistemic and processing properties.

It is now well established that cross-modal effects are common: the processing carried out by a single perceptual system is affected by information garnered by more than one sensory apparatus (Spence & Driver, 2004). For example, the perceived direction of motion of sounds is affected by information presented to the eyes (Soto-Faraco, Spence, & Kingstone, 2004). Less surprisingly, parsing a stream of speech sounds into phonemes is also multi-modal, as demonstrated by the well-known McGurk effect (McGurk & MacDonald, 1976). These results put pressure on the idea of there being a parallel array of sensory systems each with a dedicated sensory apparatus delivering a proprietary kind of information. They displace a naïve picture in which the information presented to the senses is only weighed up and integrated in the course of forming beliefs. However, the fact that some weighing and integration of information occurs at relatively early stages of processing does not on its own undermine the category of the perceptual. A single perceptual process, with apparent unity at the personal level, might take information from a number of different sensory modalities as input – it is a familiar point that the faculties of perception need not align precisely with the physical modalities that collect sensory input.

Other multi-modally-driven processes are less obviously taking place within a single perceptual modality, and are phenomenologically less like paradigmatic perceptual states. One set of examples is furnished by systems that detect properties that are more ‘abstract’ than low-level perceptual properties like colour and shape, pitch and timbre. We see an object as a dog, recognise a blob as a face, identify the presence of a particular friend by their voice, see a pattern of motion as biological, and so on. Many of these ways of detecting distal properties can be driven by several different sources of sensory input. These processes may not be entirely insulated against the influence of background information, but nor are they a matter of coming to considered judgements on the basis of everything we believe and perceive. The intense debates in philosophy about whether such properties are properly considered perceptual or cognitive (Hawley & Macpherson, 2011) are not enough to show that there is no fact of the matter about the question. Nor does the existence of borderline cases show that the category of the perceptual can do no useful work. But one possible diagnosis of the difficulty of drawing a line that corrals off the purely perceptual is that any putative way delineating the perceptual fails to capture a class that has sufficiently distinctive properties to play a special explanatory role. Questions that are typically asked about paradigmatically perceptual cases may be answered in similar ways when asked of other kinds of psychological processing.

An even greater challenge is presented by systems that are input-driven but have an amodal phenomenology. An example is the capacity to represent one’s own spatial location. People and other animals represent the relative locations of various features of their environment, and their own location within that map (Gallistel, 1990). That capacity probably depends in part on an evolutionarily ancient system that is shared at least with rodents, dependent in part on the firing of place cells in the hippocampus (O’Keefe & Burgess, 1996), which stores map-type information about the relative position of objects in the environment, and where the animal uses sensory cues to keep track of its current location (O’Keefe & Nadel, 1978). This system is genuinely amodal, in that it takes as input whichever kinds of sensory information are relevant in the current circumstances. Allocentric location can be updated on the basis of visual cues, olfactory cues or, in the absence of these, by the animal integrating over its current speed and trajectory. But in other respects representation of current location is like a perceptual process, in that its outputs are strongly constrained by sensory input, and are by no means an all-things-considered judgement based on everything else the animal represents.

A case where we frequently experience the relative encapsulation of an amodal input system is the parsing of grammatical structure. I can hear a garden-path sentence as ungrammatical while at the same time knowing that it is grammatical, perhaps even while knowing how it ought to be parsed (Caplan & Waters, 1999). Granted, the online parsing of a sentence is affected by background knowledge: statistical expectations based on word frequency, anticipations about likely meaning based on semantic knowledge, and other kinds of expectation like scripts and thematic relations. Parsing can be driven by spoken language, sign language or written language, or a combination of visual and aural cues, and is affected by other perceived aspects of the situation like the emotional context. So it is a process that relies on integrating a large array of different sources of information. Yet it has a fast and mandatory aspect that allows it to produce outputs at odds with our all-things-considered beliefs, which distinguishes it from paradigm cognitive states.

Susan Carey has emphasised another set of examples which she calls the systems of “core cognition” (Carey, 2009). These systems are intermediate between the paradigmatically perceptual and the paradigmatically cognitive. Two of her flagship cases concern numerosity. Carey (2009) marshals an impressive array of evidence for the existence of two different relatively low-level systems that are involved in representing quantities. The first is the object file system, which individuates small arrays of objects in parallel and keeps track of which is which as they move. While the numerosity of these sets is not represented explicitly, numerosity is implicit in the way the system operates: comparing arrays via 1-1 correspondence and keeping track of the addition or subtraction of small numbers of objects from the set. The second is the analogue magnitude system which is capable of keeping track of the approximate number of items in a large set (Dehaene, 1997).

Carey argues that these processes deserve their own category in the psychological inventory. They are neither clearly perceptual nor clearly cognitive. They operate amodally, on a variety of modal inputs, but the calculations they perform are informationally encapsulated and relatively independent of what is going on in the rest of cognition. Carey also argues that representations of agency are part of “core cognition”: when objects move in certain ways it just looks as if they are agents, driven not by external forces but by their own internal goals (think of cartoons of geometric figures moving in agentive ways (Csibra et al., 1999)).

Another example on the borderline is in the perception of causation. In Michotte-style experiments (Michotte, 1963) the precise timing of the movement of two circles on a screen can make it appear as if the first hits the second and causes it to move (“launches” it). Introduce a slight delay, and it looks as if the second circle moves off on its own (Scholl & Tremoulet, 2000). The experienced difference between the two settings is input-driven and partially encapsulated (e.g. against the knowledge that both are a matter of lights on a computer screen and neither is causal), but it is widely disputed whether a causal relation is something that can be perceived, as opposed to being contributed by cognition (Siegel, 2009).

Some might argue that detection of agency and causation are simply examples of ‘high-level’ perceptual experience. Other philosophers argue that the contents of genuinely perceptual experience are limited to more ‘low-level’ features, like colour and location, and pitch and timbre, in which case the agency and causation cases are non-perceptual. Either way, it is reasonably clear that these are not simply cases of belief. There is a contrast to be drawn between agency being represented by one of Carey’s systems of core cognition (e.g. with the animated triangles) and having a personal level belief that something is an agent. So if these cases are to be counted as non-perceptual, then the perception-cognition distinction cannot cover all the cases. More seriously, these cases suggest that the range of input systems, on which belief formation eventually relies, is much broader than the paradigmatic examples of perceptual processing suggest. So in the project of characterising how beliefs are based on and justified by the processing of information from the senses, characterising these systems should be just as important. Furthermore, their borderline nature raises the possibility that there may be no way, that does distinctive and important explanatory work, of distinguishing them from perceptual states.

A final example is less familiar. There are many experimental paradigms in which subjects are able to solve a task without having any conscious awareness of how they do it (Dienes & Perner, 2003). Sometimes when subjects learn this kind of task they have no idea at

all that they're getting it right; they report a phenomenology of guessing. In other situations subjects report a 'feeling of familiarity'. It turns out that, in those situations, the feeling of familiarity is a good guide as to whether they have learnt to perform the task accurately, but subjects still have no idea at all how they do it (Scott & Dienes, 2008). This feeling of familiarity is an amodal representation, triggered by external stimuli (the subject is familiar with some tasks but not others), and which feeds into downstream processing including verbal report and belief formation, but without itself being readily classifiable as either perceptual or cognitive.

These examples suggest that the range of input-driven psychological processes that produce representations of transitory features of the current environment extends far beyond the paradigmatically perceptual. Other ways of distinguishing the perceptual from the cognitive don't readily solve the problem, either. Perhaps perception is analogue, or obeys a 'picture principle' (that parts of the representation represent parts of the thing represented – (Fodor, 2007)). Another suggestion is that perceptual representations are iconic, in the sense that they represent in virtue of an isomorphism between representation and represented. Relatedly, it could be that perceptual representations are non-conceptual in the sense that they have no semantically-significant constituent structure that represents individuals, properties, or anything else below the level of a complete 'saturated' proposition. A third possibility is that there is a phenomenological distinction between perceptual systems and other forms of processing of input; perhaps that the genuinely perceptual representations are more phenomenologically salient. However, each of these suggestions is highly controversial as being criterial of perceptual experience.<sup>5</sup> In respect of each, it looks to be a substantive rather than straightforwardly definitional question whether perception has that feature. More importantly, none covers all the borderline cases discussed above. So if our focus is on explaining the set of input-driven psychological processes through which we gather new information about the world, none of these criteria will serve to count them all as perceptual.

The examples above put pressure on there being any clear personal level philosophical distinction between the perceptual and the cognitive that can do any deep explanatory work. Nor can we appeal to scientific psychology to vindicate the distinction. There it is not relied on as a theoretically important tool. Some psychological processes are indeed paradigmatically perceptual, but there is no clear dividing line between the perceptual and the cognitive, nor a clear continuum that plays any deep theoretical role in experimental psychology. Instead, the rough-and-ready distinction used in psychology is based on a loose assumption that the processing of incoming information takes place hierarchically, in a way that maps onto neural areas, with features that are readily extracted from incoming information being represented in primary sensory cortices, feeding forward to the processing of higher level features, and

---

<sup>5</sup> A different tactic is to suggest that only certain kinds of property are perceptible. If that claim is based in the response-dependence of those properties or other features of our actual perceptual apparatus, then it presupposes a characterisation of which psychological processes are perceptual and so would not help for our purposes of drawing a perception-cognition distinction. On the other hand, if it is not a claim made relative to our actual perceptual apparatus, the claim that some physical properties are just not amenable to being perceived is hard to credit.

eventually arriving at the more mysterious central systems of personal-level thought. On no view can that hierarchy be strictly delineated.

It could turn out that the top-down / bottom-up distinction could be used to identify a distinctively 'perceptual' way of processing sensory input, if it were to turn out that some input processes were immune from top-down influences entirely. That would be true, for example, if there were perceptual modules in the strict Fodorian sense. Then the collection of psychological processes that were immune from top-down influences might form a theoretically-important class, one to which special epistemological principles apply, say. That result would be a vindication rather than a defeat for the view about the importance of the top-down / bottom-up distinction advocated here. But it seems unlikely on the current state of the evidence. There are many putative examples of top-down effects even on relatively early processing of sensory input (Di Lollo et al., 2000; Summerfield & Koechlin, 2008; Ulzen et al., 2008). There appear to be many parallel processes at a level, connections across levels, and loops back from later to earlier processing. Even a strict hierarchy of information processing would only support a graded distinction between the more perceptual and more motoric, on the one hand, and the more cognitive, on the other. But when the interrelations between psychological processes are as richly intertwined as they have been found to be, we cannot appeal to psychology to substantiate any more than a rough-and-ready distinction based on how closely a process is tied to sensory inputs.

This rapid canter through the perception-cognition distinction is too brief to establish a positive argument against the distinction. Perhaps one of the ways of formulating it discussed above will turn out to mark a deep, explanatory-important divide. For the purposes of this paper I just want to motivate the idea that the distinction may be more problematic than commonly supposed. Since the category of the perceptual does not derive strong support from psychology or neuroscience, and may turn out not to mark a theoretically-important divide for philosophical purposes, there is merit in reformulating issues from the cognitive penetrability debate in a way that does not take for granted that the perceptual forms a theoretically important category. One merit of the distinction between top-down and bottom-up effects, as clarified in this paper, is that it allows us to do so. We can simply ask, of any psychological process: to what extent and in which respects is it driven by bottom-up information, and in which by top-down information?

## **5 Conclusion**

The distinction between top-down and bottom-up effects is relied upon widely in psychology and cognitive neuroscience. Its fruitful use in the science suggests that it corresponds to an empirically real distinction. A *prima facie* problem with the way standard psychological definitions capture the distinction can be overcome by appealing to some reasonably well-understood philosophical resources: the distinctions between occurrent and dispositional representation, implicit and explicit representation, and the contrast between more and less direct computational influences. So-clarified we can ask, of any psychological process, to what extent and in which ways it is driven by bottom-up information, and in which ways influenced by top-down information. The answer forms an important part of the story of how sensory input is made use of in psychological processing. It also forms the basis for an epistemological assessment of the output of the process. A further merit is that the top-down vs. bottom-up

distinction can be applied to any psychological process. Thus, if it turns out that the category of perceptual processes, presupposed by the literature on cognitive penetrability, cannot be individuated in a distinctive way that plays a theoretically important role, some of the issues in the cognitive penetrability debate can still be addressed in terms of the top-down / bottom-up distinction.

Nicholas Shea

4 November 2011

## Acknowledgements

The author would like to thank Tim Bayne, Martin Davies, Zoltan Dienes and James Stazicker for discussion and comments on earlier drafts; the audience at 'At the interface between perception and cognition' in Oxford for helpful discussion; and Dustin Stokes, Mohan Matthen and Stephen Biggs for helpful comments on a previous draft. This work was supported by: the Wellcome Trust (grant number 086041), the Oxford Martin School and the John Fell OUP Research Fund.

## References

- Caplan, D., & Waters, G. S. (1999). Verbal working memory capacity and language comprehension. *Behavioral and Brain Science*, 22, 114–126.
- Carey, S. (2009). *The Origin of Concepts*. Oxford: O.U.P.
- Carlson, N. R., Miller, H., Heth, C. D., Donahoe, J. W., & Martin, G. N. (2010). *Psychology: the science of behavior* (7th (International) ed.). Boston, MA: Allyn & Bacon (Pearson).
- Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: the perception of 'pure reason' in infancy. *Cognition*, 72(3), 237–267.
- D'Angelo, E., & De Zeeuw, C. I. (2009). Timing and plasticity in the cerebellum: focus on the granular layer. *Trends in neurosciences*, 32(1), 30–40.
- De Zeeuw, C. I., Hoebeek, F. E., Bosman, L. W. J., Schonewille, M., Witter, L., & Koekkoek, S. K. (2011). Spatiotemporal firing patterns in the cerebellum. *Nature Reviews Neuroscience*.
- Dehaene, S. (1997). *The Number Sense*. Oxford: OUP.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *Journal of Experimental psychology: General*, 129(4), 481–507.
- Dienes, Z., & Perner, J. (2003). Unifying consciousness with explicit knowledge. In A. Cleeremans (Ed.), *The unity of consciousness: Binding, integration, and dissociation* (pp. 214–232). Oxford OUP.
- Eliasmith, C., & Anderson, C. H. (2003). *Neural Engineering: computation, representation, and dynamics in neurobiological systems*. London / Cambridge MA: MIT Press.
- Eysenck, M. W. (1998). *Psychology: an integrated approach*. Harlow: Addison Wesley Longman.
- Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. *Vision Research*, 35, 3003–3013.
- Fahle, M., & Poggio, T. (2002). *Perceptual Learning*. Cambridge, MA: MIT Press.
- Fiorentini, A., & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature*, 287, 43–44.

- Fodor, J. (2007). The revenge of the given. In B. P. McLaughlin & J. Cohen (Eds.), *Contemporary debates in philosophy of mind* (pp. 105–116). Oxford, U.K.: Blackwell.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127-138.
- Friston, K., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, 159(3), 417-458.
- Gallistel, C. R. (1990). *The Organization of Learning*. London / Cambridge MA: MIT Press.
- Hawley, K., & Macpherson, F. (Eds.). (2011). *The Admissible Contents of Experience*: Wiley-Blackwell.
- Kohn, A. (2007). Visual Adaptation: Physiology, Mechanisms, and Functional Benefits. *Journal of neurophysiology*, 97, 3155-3164.
- Lyons, J. (forthcoming). Circularity, Reliability, and the Cognitive Penetrability of Perception. *Philosophical Issues*.
- Macpherson, F. (2011). Cognitive Penetration of Colour Experience: Rethinking the Issue in Light of an Indirect Mechanism. *Philosophy and Phenomenological Research*.
- Maertens, M., & Pollmann, S. (2005). fMRI Reveals a Common Neural Substrate of Illusory and Real Contours in V1 after Perceptual Learning. *Journal of Cognitive Neuroscience*, 17(10), 1553-1564.
- Mazzoni, P., & Krakauer, J. W. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. *The Journal of neuroscience*, 26(14), 3642.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices.
- Mechelli, A., Price, C. J., Noppeney, U., & Friston, K. J. (2003). A dynamic causal modeling study on category effects: Bottom-up or top-down mediation? *Journal of Cognitive Neuroscience*, 15(7), 925-934.
- Michotte, A. (1963). *The perception of causality*. Oxford: Basic Books.
- O'Keefe, J., & Burgess, N. (1996). Geometric determinants of the place fields of hippocampal neurons. *Nature*, 381(6581), 425-428.
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Clarendon Press.
- Poggio, T., Fahle, M., & Edelman, S. (1992). Fast Perceptual Learning in Visual Hyperacuity. *Science*, 256, 1018-1021.
- Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in cognitive sciences*, 4(8), 299-309.
- Scott, R. B., & Dienes, Z. (2008). The conscious, the unconscious, and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(5), 1264.
- Siegel, S. (2009). The visual experience of causation. *The Philosophical Quarterly*, 59(236), 519-540.
- Siegel, S. (2011). Cognitive Penetrability and Perceptual Justification. *Noûs*.
- Soto-Faraco, S., Spence, C., & Kingstone, A. (2004). Cross-Modal Dynamic Capture: Congruency Effects in the Perception of Motion Across Sensory Modalities. *Journal of Experimental Psychology: Human Perception and Performance*, 30(2), 330.
- Spence, C., & Driver, J. (2004). *Crossmodal space and crossmodal attention*: Oxford University Press, USA.
- Stokes, D. (2011). Perceiving and desiring: a new look at the cognitive penetrability of experience. *Philosophical Studies*.
- Summerfield, C., & Koechlin, E. (2008). A Neural Representation of Prior Information during Perceptual Inference. *Neuron*, 59, 336-347.
- Ulzen, N. R. v., Semin, G. R., Oudejans, R. R. D., & Beek, P. J. (2008). Affective stimulus properties influence size perception and the Ebbinghaus illusion. *Psychological Research*, 72, 304-310.
- Watanabe, T., Naney, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature*, 413, 844-848.