Multimodal mental imagery and perceptual justification

There has been a lot of discussion about how the cognitive penetrability of perception may or may not have important implications for understanding perceptual justification. The aim of this paper is to argue that a different set of findings in perceptual psychology poses an even more serious challenge to the very idea of perceptual justification. These findings are about the importance of perceptual processing that is not driven by corresponding sensory stimulation in the relevant sense modality (such as amodal completion and multimodal completion). I argue that these findings show that everyday perception is in fact a mixture of sensory stimulation-driven perceptual processing and perceptual processing that is not driven by corresponding sensory stimulation in the relevant sense modality and we have strong reasons to doubt the epistemic pedigree of the latter process. The implication of this is not that we should become skeptics or deny the possibility of perceptual justification. It is, rather, that the only way in which we can understand when and whether a perceptual state justifies beliefs is by paying close attention to empirical facts about the reliability of perceptual processing that is not driven by corresponding sensory stimulation in the relevant sense modality. In this sense (a very narrow sense) epistemology needs to be naturalized.

I. Two empirical problems with perceptual justification

Some perceptual states can justify beliefs. If I look out of the window and see that it is raining outside, this may (and as a default, would) justify my belief that it is raining outside. But this simple picture of the relation between perception and belief has been questioned recently. A vast number of experiments in perceptual psychology purport to show that perception is not an encapsulated process: it can be influenced by our beliefs. And these findings about top-down influences on perception are said to have important implications for the potential epistemic role perception may play. The general line of argument is this: if perception is cognitively penetrated, then it is not an unbiased way of learning about the world as our preexisting thoughts, beliefs and expectations will influence how and what we perceive.

So we get a vicious circularity: our beliefs, thoughts and expectations are supposed to be based on and justified by our perceptual states, but these perceptual states themselves are influenced by our beliefs, thoughts and expectations (because of cognitive penetration). What we get is some kind of ‘wishful seeing’ (see Pylyshyn 1999, who argues that seeing is not wishful in this sense). As Roberto Bolano says in the novel *2666*, “People see what they want to see and what people want to see never has anything to do with the truth”.[[1]](#footnote-1)

 The challenge from cognitive penetration was originally focusing on one specific account of perceptual justification, namely, dogmatism (Siegel 2011, see also Lyons 2011): the view that  "whenever you have an experience as of p, you thereby have immediate prima facie justification for believing p" (Pryor 2000, p. 536). The argument was that if perception is cognitively penetrated, dogmatism is not an option because the perceptual states that our beliefs are supposed to be justified by are themselves influenced by our existing beliefs and expectations. This argument has been generalized to apply to other theories of justification (not just dogmatism, see Siegel 2011, Tucker 2014, see also Lyons 2015, Ghijsen 2016, Silins 2016).

The main aim of this paper is to argue that even if the cognitive penetrability of perception is not something theories of perceptual justification need to worry about, a different set of findings in perceptual psychology poses a much more serious challenge to the very idea of perceptual justification. These findings are about the importance of perceptual processing that is not driven by corresponding sensory stimulation in the relevant sense modality (such as amodal completion and multimodal completion).

 I argue that these findings show that everyday perception is in fact a mixture of sensory stimulation-driven perceptual processing and perceptual processing that is not driven by corresponding sensory stimulation in the relevant sense modality and we have strong reasons to doubt the epistemic pedigree of the latter process. The implication of this is not that we should become skeptics or deny the possibility of perceptual justification. It is, rather, that the only way in which we can understand when and whether a perceptual state justifies beliefs is by paying close attention to empirical facts about the reliability of perceptual processing that is not driven by corresponding sensory stimulation in the relevant sense modality. In this sense (a very narrow sense) epistemology needs to be naturalized.

II. The first problem: Top-down influences on perception

One influential debate about perception is about whether it is an encapsulated process that is protected from any kind of top-down influences or it is influenced and modified by top-down information. What complicates this debate, often referred to as the cognitive penetrability debate, is that it is not at all clear what kind of mental state is supposed to be doing the penetrating and what kind of mental state is supposed to be penetrated. In other words, it is not clear what is ‘top’ and what is ‘below’ in the debate about top-down influences on perception.

 Once we clarify these conceptual issues, it seems that there is a wealth of empirical evidence in favor of the claim that there are indeed some top-down influences on perception. But then the question becomes: is this something epistemologists would have to worry about? I will argue that they wouldn’t.

The main conceptual confusion concerning debates about top-down influences on perception is that it is not clear what is meant by ‘perception’ in this context. Some (especially philosophers (Siegel 2011, Macpherson 2012, Stokes 2012), but also psychologists, see, e.g., Firestone and Scholl 2014, 2016) take ‘perception’ in this context to be perceptual experience: something we are consciously aware of. If we work with this concept of perception, then the question is whether top-down influences can alter the way we experience a scene – the phenomenal character of our experience: what it is like to perceive this scene.

Another way of understanding what is meant by ‘perception’ when we talk about top-down influences on perception is perceptual processing – something neuroscientists (and also some psychologists) worry about. Here the question is whether processing in, say, the primary visual cortex is influenced in a top-down manner.

These two questions are clearly very different – one of them is about phenomenology and the other is about early perceptual processing. And as changes in early perceptual processing are neither necessary nor sufficient for changes in perceptual phenomenology, there is no easy traffic between these two different sub-debates.

The methodology of addressing these questions is also different. When it comes to figuring out what influences our perceptual phenomenology, we need to rely on self-reports and introspection (together with some form of inference to the best explanation). When it comes to early perceptual processing, we need to rely on neuro-imaging data (and perhaps also some behavioral data, like reaction-time). The problem is that it is very rarely made clear which of these very different questions are being asked.

I argued elsewhere (Teufel and Nanay 2017) that it is doubtful that the first of these debates could ever be resolved in a satisfactory manner. One reason for this is the well-documented unreliability of introspection (see Schwitzgebel 2008, Spener and Bayne 2010) – whereas introspection is something this debate needs to heavily rely on. But an even more important worry is that any way of resolving this debate would need to appeal to a clear and unproblematic way of keeping apart perceptual and non-perceptual phenomenology – something that we do not have. Given that it is not clear what perceptual phenomenology is and how to keep it apart, introspectively, from non-perceptual phenomenology, the question about perceptual phenomenology (of whether perceptual phenomenology depends on top-down influences) inherits all these problems. It is unclear then how we can make any progress in answering this question – if we take the question about top-down influences to be about perceptual phenomenology.

 So, as a result, I take it that a more promising (or more productive) debate about top-down influences on perception is about whether early perceptual processing is influenced in a top-down manner. And here we find a lot of extremely convincing neuroimaging and behavioral evidence for the existence of top-down influences (see, for example, Murray et al. 2002, Gandhi et al. 1999, O’Connor et al. 2002). But then the question becomes: are these results relevant from the point of view of epistemology?

Observant readers may have noticed that I talked about ‘top-down influences on perception’ and not about cognitive penetration so far. There is a reason for this. The term ‘cognitive penetration’ suggests that whatever is doing the penetration is a cognitive state and this is not something I want to be built into the very notion I am analyzing.

 When I talk about ‘top-down’ influences on perception, I want to allow for any ‘top-down’ influences – not just those that are labeled ‘cognitive’. And it is not very clear why the label ‘cognitive’ is singled out. ‘Cognitive’ can mean many things. It is sometimes contrasted with ‘affective’, but this is clearly not something we want to do if we are interested in top-down influences on perception as there may be affective influences on perception and they may be as important as (or more important than) non-affective cognitive influences (Schupp et al. 2004, Schmitz et al. 2009, Pessoa and Ungerleider 2005). The term ‘cognitive’ is also often contrasted with ‘conative’, but this also a dubious usage in the present context as there may be very good reasons to posit top-down influences on perception where it is a desire or an intention that influences our perceptual processing (Nanay 2006, Stokes 2012).

 Of course, the most straightforward use of ‘cognitive’ may just be one where it is contrasted with ‘perceptual’, but this simplifies things considerably. In fact, one reason why it is better to focus on the debate about whether there are top-down influences on early perceptual processing than on the one about whether there are top-down influences on perceptual phenomenology is that if we focus on the latter debate, the only kind of top-down influence we can talk about is from mental states (normally with non-perceptual phenomenology) to mental states with perceptual phenomenology (ie, perceptual experiences). But we have seen that addressing any questions about the presence or absence of such top-down influences then requires a very clear distinction between perceptual and non-perceptual phenomenology and we don’t have any distinction, let alone a clear one.

 If, on the other hand, we consider the debate about whether there are top-down influences on early perceptual processing, we get a more detailed picture. We have a very clear idea of the sequence in which ‘bottom-up’ perceptual processing proceeds. To take the visual sense modality as an example (Katzner and Weigelt 2013, Grill-Spector and Malach 2004, Van Essen 2004, Bullier 2004), in humans and nonhuman primates, the main visual pathway connects neural networks in the retina to the primary visual cortex (V1) via the lateral geniculate nucleus (LGN) in the thalamus; outputs from V1 activate other parts of the visual cortex and are also fed forward to a range of extrastriate areas (V2, V3, V4/V8, V3a, V5/MT). So visual processing goes from the retina via the LGN to the primary visual cortex (V1), and then to the secondary visual cortex (V2) and then (in the case of color vision) to the V4, and so on. So we have a natural ordering of mental processes in terms of what is bottom and what is top. If there are influences from V4 to the primary visual cortex, it is a top-down influence because it is from a mental process (V4) that comes later in visual processing than the primary visual cortex (V1), which is the mental process that is being influenced.

 If we raise the question of top-down influences on perceptual processing, we get a more nuanced picture. The question of top-down influences is no longer a yes or no question as in the case of the phenomenology interpretation (either there is cognitive penetration or there isn’t), but a multifaceted one. Maybe the primary visual cortex is influenced in a top-down manner by V2 and V4, but not by our expectations and beliefs. Or maybe it is only influenced by V2. Or maybe also by our expectations and beliefs. All of these claims would assert top-down influences on early perceptual processing (of which, as we have seen, there is very strong evidence), but it matters a lot what kind of top-down influences they are (see Nanay and Teufel 2016 for a detailed analysis of various kinds of top-down influences on early cortical perceptual processing and the differences between those top-down influences that come from within the visual system and those that come from post-perceptual processing).

While some of these findings show very rigorously that there are top-down influences on perception, it is important to see that many of these findings are irrelevant in the present context. Epistemologists are worried about whether our perception is influenced by our beliefs or other cognitive states. So top-down influences on perceptual processing would be relevant if the ‘top’ in these influences were beliefs or other cognitive states. That is why epistemologists are interested in cognitive penetration as opposed to any form of top-down influences (where I take cognitive penetration to be the penetration of perception by beliefs or other cognitive states – thus the label ‘cognitive’). Epistemologists are not at all interested in whether the primary visual cortex is influenced by the V4/V8 or the STS or the MT.

But don’t we have empirical evidence that perceptual processing is influenced by non-perceptual or cognitive states? While some (including myself, see Teufel and Nanay 2016) would say so, this premise would be easy to dispute. Any such argument would need to presuppose some form of firm distinction between perceptual and non-perceptual processing (not between perceptual and non-perceptual phenomenology as before, but perceptual and non-perceptual processing) – so that we can zero in on those top-down influences that go across this divide. The problem is that, depending on how we draw this divide between perceptual and non-perceptual, any finding about top-down influences on perception could be framed as merely showing an intra-perceptual top-down effect.

 To sum up, there are two questions of top-down influences on perception. One (concerning perceptual processing) is by and large irrelevant for epistemology. The other (about perceptual phenomenology) would be relevant for epistemology but it is difficult to see how this debate could be resolved in a satisfactory manner. As a result, top-down influences on perception (or cognitive penetration) should not give a significant cause for concern for anyone who is interested in perceptual justification. Empirical findings about top-down influences on perception do not seem to jeopardize any philosophical account of perceptual justification.[[2]](#footnote-2)

 But there is another set of empirically findings about perception that should worry anyone who is interested in perceptual justification much more: the findings about the importance of perceptual processes that are not triggered by corresponding sensory stimulation in the given sense modality.

III. The second problem: imagery-infused perception

Some perceptual processing starts with sensory stimulation. The light hits our retina and vision is the complex visual processing of this sensory stimulation. This perceptual processing may include, depending on whom you ask, the interpretation or the elaboration or the embellishment of the sensory stimulation, but it is the sensory stimulation that is processed/interpreted/elaborated on.

 But some other cases of perceptual processing are not the processing of sensory stimulation because there is no sensory stimulation to be processed. These perceptual processes are not triggered by corresponding sensory stimulation in the relevant sense modality.

 In psychology and neuroscience, perceptual processes that are not triggered by corresponding sensory stimulation in the relevant sense modality are called ‘mental imagery’. Here is a representative quote from a recent review paper: “We use the term ‘mental imagery’ to refer to representations […] of sensory information without a direct external stimulus” (Pearson et al. 2015). I realize that this way of talking about mental imagery may be somewhat controversial for philosophers, but I will use the concept of mental imagery as a technical concept in this paper and use the term ‘mental imagery’ as a shorthand for ‘perceptual processing that is not triggered by corresponding sensory stimulation in the relevant sense modality’. Those readers who *really* dislike this way of using the concept of mental imagery could just substitute ‘mental imagery\*’ (or ‘phantom perception’ or ‘offline perception’ or whatever they like) for ‘mental imagery’.

I will argue that we have substantial empirical evidence that the vast majority of what we take to be perception is really a mixture of sensory stimulation-driven perceptual processing and perceptual processing that is not triggered by corresponding sensory stimulation in the relevant sense modality. In other words, we have substantial empirical evidence that the vast majority of what we take to be perception is really a mixture of sensory stimulation-driven perceptual processing and mental imagery. And these findings pose a serious challenge to the very idea of perceptual justification.

Before we go on, I should emphasize that ‘perceptual processing’ in this definition means early cortical perceptual processing (Katzner and Weigelt 2013, Grill-Spector and Malach 2004, Van Essen 2004, Bullier 2004). In the case of sensory stimulation-driven perception, we have a correspondence between this perceptual processing and the sensory stimulation. In the visual sense modality, for example, early perceptual processing is retinotopic. The primary visual cortex (and also many other parts of the visual cortex see Grill-Spector and Malach 2004 for a summary) is organized spatially in a way that is very similar to the retina – it is retinotopic. While this retinotopy of the early visual cortices (and their equivalent in the other sense modalities, see, e. g., Talavage et al. 2004) is an extremely convenient way of gaining evidence about the correspondence or lack thereof of sensory stimulation and perceptual processing, this is just one way in which the two can correspond. There are others. In other words, the correspondence between sensory stimulation and perceptual processing does not have to be retinotopic.[[3]](#footnote-3)

Finally, visual mental imagery is *visual* perceptual processing that is not triggered by *visual* sensory stimulation. As we shall see below, visual mental imagery can be (and is often) triggered by *non-visual* (for example, auditory) sensory stimulation. I call this form of mental imagery multimodal mental imagery.

IV. Varieties of mental imagery

I’ve defined mental imagery as perceptual processing that is not triggered by corresponding sensory stimulation in the relevant sense modality. It is important that this definition remains silent on a number of features of mental imagery. For example it allows for voluntary and involuntary mental imagery (as perceptual processing that is not triggered by corresponding sensory stimulation in the relevant sense modality may or may not be voluntary). It allows for conscious and unconscious mental imagery (as there is no restriction that perceptual processing that is not triggered by corresponding sensory stimulation in the relevant sense modality would need to be conscious). And it is also neutral about whether mental imagery is accompanied by the feeling of presence.

As I indicated above, if the reader has a strong conviction that mental imagery is necessarily conscious (or necessarily voluntary or necessarily not accompanied by the feeling of presence), she can take the rest of the discussion to be about a very different technical concept, ‘mental imagery\*’. I defend the plausibility of this way of thinking about mental imagery (and, as a result, the continuity between the psychological/neuroscientific and the philosophical/everyday conception of mental imagery) elsewhere (Nanay forthcoming a, forthcoming b), but I do not need it for the argument in this paper.

Many different kinds of perceptual processes will count as mental imagery according to this definition, some of which may not be normally categorized as such. Here are a couple of examples:

The blind spot of the retina cannot be stimulated – there are no receptors there. If the light hits this part of the retina it gives rise to no sensory processing. So we receive no sensory information from that region of the retina. Nonetheless, our perceptual system ‘fills in’ the sensory input of the blind spot on the basis of the sensory input of the surrounding parts of the retina. The perceptual processing of information at the blind spot region of the visual field happens already in early visual cortices (Komatsu et al. 2000, Ramachandran 1992, Awater et al. 2005, Spillman et al. 2006, Fiorani et al. 1992), but it is not triggered by corresponding sensory stimulation because there is no sensory stimulation at the blind spot, let alone corresponding sensory stimulation.[[4]](#footnote-4) Processing at the blind spot counts as mental imagery.

As does peripheral vision. Peripheral regions of the retina are much less sensitive than focal ones. And this focal preference is even stronger in early cortical processing. As a result, the properties of the peripheral regions of the visual field that our perceptual system processes are much less determinate than the properties of the focal regions. This asymmetry is especially striking when it comes to color vision as there are very few retinal cells in the periphery that are sensitive to color information (Hansen et al. 2009). But the same is true for all other perceptually processed properties, like size or shape. Peripheral vision can also ‘fill in’ some regions of the periphery. ‘Artificial scotoma’ is a region of the visual field where different sensory stimulation is induced from what surrounds it (and this can be no sensory stimulation surrounded by a pattern). If this is presented in the periphery, the visual system fills in the scotoma, making it blend in. This filling in process starts very early in the visual processing (De Weerd et al. 1995, 1998, 2006, Ramachandran and Gregory 1991, Weil et al 2007, 2008, Troncoso et al. 2008, Welchman and Harris 2001). Again, we have perceptual processing that is not triggered by corresponding sensory stimulation (as the sensory stimulation at the artificial scotoma is very different from what is perceptually processed).

Amodal completion also counts as mental imagery according to my definition. Amodal completion is the representation of those parts of the perceived object that we get no sensory stimulation from. In the case of vision, it is the representation of occluded parts of objects we see: when we see a cat behind a picket fence, our perceptual system represents those parts of the cat that are occluded by the picket fence. In tactile perception, it is the completion of those parts of the objects we touch that are not in direct contact with our hand, for example. We complete those parts amodally.[[5]](#footnote-5) Amodal completion is, by definition, perceptual processing that is not triggered by corresponding sensory stimulation. The mental imagery involved in amodal completion may bring about very similar issues for the very idea of perceptual justification as the ones I raise about multimodal mental imagery in this paper – see Helton and Nanay in preparation).

 But this paper is about a form of mental imagery where perceptual processing in one sense modality is triggered by sensory stimulation in another sense modality: this paper is about multimodal mental imagery.

V. Multimodality and mental imagery

Multimodal perception is the norm and not the exception – our sense modalities interact in a variety of ways (see Spence & Driver 2004, Vroomen et al. 2001, Bertelson and de Gelder 2004 for summaries and O’Callaghan 2008a, 2011 as well as Macpherson 2011 for philosophical overviews). Information in one sense modality can influence and even initiate information processing in another sense modality at a very early stage of perceptual processing (even in the primary visual cortex in the case of vision, for example, see Watkins et al. 2006).

A simple example is ventriloquism (Bertelson 1999, O’Callaghan 2008b). The auditory sense modality identifies the ventriloquist as the source of the voices, while the visual sense modality identifies the dummy. And as a result of the influence of vision, we auditorily experience the voices as coming from the dummy.

What I am interested in here is not multimodal perception, but multimodal mental imagery: cases where there is perceptual processing in one sense modality that is not triggered by corresponding sensory simulation in that sense modality, but rather by corresponding sensory stimulation in a different sense modality.

Here is an example from Nanay forthcoming a: When I am looking at my coffee machine that makes funny noises, this is an instance of multisensory perception – I perceive this event by means of both vision and audition. But very often we only receive sensory stimulation from a multisensory event by means of one sense modality. If I hear the noisy coffee machine in the next room, that is, without seeing it, then the question arises: how do I represent the visual aspects of this multisensory event?

We know that our visual system in these circumstances does get activated (and even the very early visual cortical areas can, see Hertrich et al. 2011, Pekkola et al. 2005, Zangaladze et al. 1999, Ghazanfar & Schroeder 2006, Martuzzi et al. 2007, Calvert et al. 1997, James et al. 2002, Chan et al. 2014, Hirst et al. 2012, Iurilli et al. 2012, Kilintari et al. 2011, Muckli & Petro 2013, Vetter et al. 2014). In other words, there is early cortical activation in the visual sense modality without corresponding sensory stimulation in this sense modality. That is, we represent these features by means of mental imagery. I call this form of mental imagery multimodal mental imagery.

Remember the last phrase in the definition of mental imagery: perceptual processing that is not triggered by corresponding sensory stimulation *in the relevant sense modality*. This phrase is crucial in the present context. Mental imagery can be triggered by corresponding sensory stimulation as long as it is not in the relevant sense modality.[[6]](#footnote-6)

In other words, depending on what perceptual processing is triggered by, we get different perceptual phenomena. If perceptual processing is triggered by corresponding sensory stimulation in the relevant sense modality, we get perception. If it is triggered by corresponding sensory stimulation in another sense modality, we get multimodal mental imagery. If it is triggered by something else, we get some other kind of (non-multimodal) mental imagery. In short, multimodal mental imagery is mental imagery in one sense modality induced by sensory stimulation in another sense modality. And, as we have seen, we have strong empirical evidence that non-sensorily driven perceptual processing in any sense modality can be induced by sensory stimulation in any other sense modality.

But the vast majority of the entities we encounter are multisensory entities (that is, entities that we could encounter by means of more than one sense modalities). And our perceptual access to these multisensory entities is very rarely absolute (that is, encompassing all relevant sense modalities). If we put together these two claims, what we get is that multimodal mental imagery is the norm, not the exception.

 Here is a nice experimental illustration of this point: the double flash illusion. Subjects are presented with one flash and two beeps simultaneously (Shams et al. 2000). So the sensory stimulation in the visual sense modality is one flash. But they experience two flashes and already in the primary visual cortex, two flashes are processed (Watkins et al. 2006). This means that the double flash illusion is really about multimodal mental imagery: we have perceptual processing in the visual sense modality (again, already in V1) that is not triggered by corresponding sensory stimulation in the visual sense modality (but by corresponding sensory stimulation in the auditory sense modality).

 The picture we ended up with is one where perceptual processes consist of a sensory-stimulation-driven and a non-sensory-stimulation-driven component (where by sensory-stimulation-driven, I mean driven by corresponding sensory stimulation of the relevant sense modality). In the vast majority of perceptual scenarios, sensory stimulus-driven perceptual processing is combined with mental imagery. And in these cases, much of what we take ourselves to perceive we really (at least partly) represent by means of mental imagery.

VI. Back to perceptual justification (and top-down influences)

I argued in Section II that top-down influences on perception should not be considered to be a very strong reason to reevaluate how we think about perceptual justification. But the empirical results I introduced in Section III-V should. The primary reason why mental imagery is important in the discussion of perceptual justification is *not* that it can be influenced in a top-down manner. In this section, I argue that mental imagery can be influenced in a top-down manner – but, again, this is not the primary reason why they should pose a problem for any account of perceptual justification.

Some cases of mental imagery are fully bottom-up, but some others are influenced or even initiated in a top-down manner. There are many forms of mental imagery that do not depend on any top-down information. The kind of mental imagery our perceptual system uses to fill in the blind spot is one clear example. The perceptual processing of information that would correspond to the blind spot is not triggered by corresponding sensory stimulation in the given sense modality because there is no corresponding sensory stimulation in the given sense modality: the blindspot has no receptors. But this perceptual processing is determined fully by bottom-up information about the sensory stimulation of the parts of the retina that surround the blind spot. No top-down influence is needed and we have no evidence that there are any top-down influences on this form of mental imagery. Another fully bottom-up example of multimodal mental imagery is the double flash illusion.

But some other cases where multimodal mental imagery. One widely used and researched example of what I call multimodal mental imagery is seeing someone talking on television with the sound muted. The visual perception of the talking head in the visual sense modality leads to an auditory mental imagery in the auditory sense modality (e.g., Calvert et al., 1997; Hertrich, Dietrich, & Ackermann, 2011; Pekkola et al., 2005, see also Spence and Deroy 2013 for a philosophical summary).

The auditory mental imagery in this case clearly depends on bottom-up factors like the lip movements of the person on the screen. But not only these. If this person is someone you know or have heard speak, your auditory mental imagery will be influenced by this information. If it is Barack Obama (someone you have, presumably, heard before), you will ‘hear’ him speaking with his distinctive tone of voice or intonation, for example. This is a case (and cases like this are not at all rare) where multimodal mental imagery is influenced by top-down information.

If we put together these findings with the claim about the importance of mental multimodal imagery in everyday perception, we get a straightforward argument for top-down influences on perception: multimodal mental imagery is influenced by top-down information and perception is very much influenced by multimodal mental imagery. By transitivity, it would follow that perception is influenced by top-down information.

Note that while this argument may establish top-down influences on perception, it provides no arguments for or against the dependence of perception (either perceptual phenomenology or perceptual processing) on *cognitive* states – which would be the sense of top-down influences on perception that is relevant in the context of perceptual justification. From the point of view of epistemology and of the question of perceptual justification, the only debate about top-down influences on perception that matters is about cognitive influences on perception. Whether there are some forms of top-down influences on early perceptual processes is strictly irrelevant. The reason why the importance of multimodal mental imagery in everyday perception is important from the point of view of epistemology lies beyond the focus on top-down influences.

VIII. A deeper problem about perceptual justification

Findings about mental imagery show that much of what we take ourselves to perceive we really represent by means of mental imagery. Perceptual states are almost always of multisensory entities and unacquainted parts of these multisensory entities are represented by means of mental imagery. In the world of multimodal perception, what we take ourselves to perceive is partly represented by multimodal mental imagery. But then we have much stronger reasons to think that perception is not the right kind of mental state to base our beliefs on and to justify our beliefs by.

Mental imagery may or may not be influenced by top-down information. And even when it is, it is not clear how far up this top-down information comes from. But this is cold comfort for those who expect perception to serve as some unbiased basis for perceptual justification. Even in those cases where mental imagery is not at all influenced by top-down information, it fails to be caused by what it represents.

Even in the somewhat trivial case of the blind spot, where, supposedly, no top-down information is being used, the blind spot is filled in by mental imagery – by perceptual processes not triggered by corresponding sensory stimulation. So no matter what way the blind spot is filled in, that has no causal connection with whatever is in front of that part of the retina. Mental imagery, by definition, fails to be caused by what it is about.

Sensory stimulation-driven perception is caused by what it is of. Light from the perceived object hits our retina and that is the sensory stimulation that gets processed. But mental imagery is by definition not sensory stimulation-driven. So it is cut off from the object it is about. One link from the causal chain is missing. Light from the perceived object hits our retina, the perceptual processing in the visual cortices is not triggered by this retinal stimulation. So the perceptual state that is supposed to do the epistemic heavy-lifting partly depends on (or maybe even constituted by, see Nanay forthcoming a, forthcoming b) mental imagery – something not particularly well-suited at all for any epistemic role.

Why not? Because at least on the face of it, it violates both the safety and the sensitivity conditions of justification. The sensitivity condition is clearly violated: the beliefs we form on the basis of multimodal mental imagery are not sensitive to possible changes in the multimodally completed bits: no change in the multimodally completed part of the multisensory event would show up in your beliefs.

The safety condition is a bit more complicated (Sosa 1999). Here is one version of the safety condition: in those close possible worlds where one believes that p on the basis of multimodal mental imagery, p is the case. A lot will depend on how one construes the scope of these possible worlds (all the closest possible worlds or most of the closest possible worlds). But even if we go with the least restrictive characterization, there will be cases, for example, the double-flash illusion I described above, where it will not be the case that in most close possible worlds where I believe p on the basis of my multimodal mental imagery, p is true. I will believe in most, but maybe even all close possible worlds that there were two flashes. But there is only one flash. This is an extreme case, but the point is that we can’t exclude the possibility of multimodal mental imagery violating the safety condition in any scenario where multimodal mental imagery plays a role (which means, as we have seen, in the vast majority of everyday perceptual scenarios). In short, almost all instances of perception are mixtures of a state that is supposed to track the truth (perception) and a state that, on the face of it, isn’t (mental imagery).

Perception is supposed to be a good way of justifying our beliefs because perception tracks truth. But mental imagery is, by definition, a step removed from the truth it is supposed to track. Of course it can track truth albeit in a fallible manner. The mental imagery used for filling in the blind spot, for example, is really very reliable. It can be fooled, but in the vast majority of cases it isn’t. So the mental imagery that is used to fill in the blind spot does track truth – not 100% reliably, but nonetheless reliably enough. And the reason we know this is that we know the exact mechanisms of how the visual system uses the sensory stimulation around the blind spot as an input when filling in the blind spot. If this mechanism were less reliable, this mental imagery would fail to track the truth.

But then the same question needs to be asked about those forms of mental imagery that play a more important role in everyday perception: about whether the mechanisms that construct these forms of mental imagery are reliable enough. And here it seems safe to say that the more bottom-up the mental imagery is (like in the blind spot case), the more reliable we should consider it to be. Nonetheless, whether perception can justify beliefs depends on empirical facts about the reliability of the mechanisms of mental imagery involved in perception.[[7]](#footnote-7)

 Again, we can make the default assumption that our perceptual system built pretty good mechanisms for mental imagery on the basis of contextual or crossmodal information that does co-vary with the scene in front of us. The blind spot is a good example. We can fool the filling in of the blind spot, but this happens very rarely and only in exceptional circumstances (and only in monocular vision, for a start).

But if what we take to be perception is really a mixture between sensory-stimulation-driven perception and mental imagery, then we cannot take it for granted that perceptual justification is unproblematic. We need to examine the mechanisms of mental imagery to see how reliable they are and what role they can play in perceptual justification.

IX. Conclusion: Epistemology naturalized

A lot more work needs to be done in order to show that we are justified to move from (imagery-infused) perception to belief. Again, this is not to say that we can’t eventually do so, we surely can. But any such move would need to involve a close empirical examination of the reliability of the processes that constitute mental imagery.

It is also important to stress that when I say that almost all of our visual perceptual states are in fact mixed sensory stimulation-driven/mental imagery states, I do not mean to suggest that the contribution of sensory stimulation-driven processes and not sensory stimulation-driven processes (that is, mental imagery) is approximately equal. In fact, it happens very rarely that they are equal. Normally, the mental imagery component is negligible. But the very fact that it is always lurking in the background should prevent us from taking perception at face value when it comes to perceptual justification.

The conclusion is that the question of perceptual justification is at least in part an empirical question – it requires the examination of the reliability of the forms of mental imagery that play a role in perception per se. This is a sense (a fairly narrow sense, to be sure) in which epistemology needs to be naturalized.

**References**

Akins, K. A. and S. Winger 1996 Ships in the night: Churchland and Ramachandran on Dennett’s theory of consciousness. In: K. A. Akins (ed.): *Perception*. pp. 173-197.

Allefeld, C. et al. (2011) Flicker-light induced visual phenomena: frequency dependence and specificity of whole percepts and percept features. *Consciousness and Cognition* 20, 1344–1362

Awater H, Kerlin JR, Evans KK, Tong F. 2005 Cortical representation of space around the blind spot. *Journal of Neurophysiology* 94: 3314-24.

Bakin, J., Nakayama, K., & Gilbert, C. (2000). Visual responses in monkey areas V1 and V2 to threedimensional surface configurations. *Journal of Neuroscience*, 20, 8188–8198.

Ban, H., H. Yamamoto, T. Hanakawa, S. Urayama, T. Aso, H. Fukuyama and Y. Ejima 2013 Topographic representation of an occluded object and the effects of spatiotemporal context in human early visual areas. *Journal of Neuroscience* 33: 16992-17007.

Bertelson, P. (1999) Ventriloquism: A case of cross-modal perceptual grouping. In: Aschersleben, G., Bachmann, T., and M¨ usseler, J. (eds): *Cognitive Contributions to the Perception of Spatial and Temporal Events*, Amsterdam: Elsevier, pp. 347–362.

Bertelson, P. and de Gelder, B. (2004) The psychology of multimodal perception. In Spence, C. and Driver, J. (eds): *Crossmodal Space and Crossmodal Attention*, Oxford: Oxford University Press, pp. 141–177.

Bullier, J. (2004). Communications between cortical areas of the visual system. In L.M. Chalupa & J.S. Werner (Eds.), *The visual neurosciences* (pp. 522–540). Cambridge, MA: MIT Press.

Bushnell, B. N., P. J. Harding, Y. Kosai and A. Pasupathy 2011 Partial occlusion modulates contour-based shape encoding in primate area V4. *Journal of Neuroscience* 31: 4012-4024.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williamns, S. C. R., McGuire, P. K., Woodruff, P. W. R., Iversen, S. D. & David, A. S. 1997 Activation of auditory cortex during silent lipreading. *Science* 276: 593-596.

Chen, J., B. Liu, B. Chen and F. Fang 2009 Time course of amodal completion in face perception. *Vision Research* 49: 752-758.

Churchland, P. S. and V. S. Ramachandran 1993 Filling in: Why Dennett is wrong. In: B. Dahlbom (ed.): Dennett and his Critics. Oxford: Blackwell, pp. 28-52.

De Weerd P, Desimone R, Ungerleider LG. 1998 Perceptual filling-in: a parametric study. *Vision Research* 38: 2721-34.

De Weerd P, Gattass R, Desimone R, Ungerleider LG. 1995 Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. *Nature* 377: 731-4.

De Weerd P, Smith E, Greenberg P. 2006 Effects of Selective Attention on Perceptual Fillingin. *Journal of Cognitive Neuroscience* 18: 335-47.

Dennett, D. C. 1991 *Consciousness Explained*. Boston, MA: Little, Brown.

Fiorani JM, Rosa MG, Gattass R, Rocha-Miranda CE. 1992 Dynamic surrounds of receptive fields in primate striate cortex: a physiological basis for perceptual completion? *Proceedings of the National Academy of Sciences U S A* 89: 8547-51.

Firestone, C., & Scholl, B. J. (2014). “Top-down” effects where none should be found: The El Greco fallacy in perception research. *Psychological Science*, *25*(1), 38–46.

Firestone, C., & Scholl, B. J. (2016). There are no top-down influences on perception. *Behavioral and Brain Sciences*, forthcoming

Gandhi, S.P. et al, (1999) “Spatial attention affects brain activity in human primary visual cortex”, *Proceedings of the National Academy of Sciences* 96 (1999): 3314-3319.

Ghazanfar, A. A. & Schroeder, C. E. 2006 Is neocortex essentially multisensory? *Trends in Cognitive Sciences* 10: 278-285.

[Ghijsen, Harmen (2016). The Real Epistemic Problem of Cognitive Penetration.](http://philpapers.org/rec/GHITRE%22%20%5Ct%20%22_blank) *Philosophical Studies*

Gregory, R. L. 1987. *The Oxford Companion to The Mind*, Oxford: Oxford University Press.

Grill-Spector, K. and R. Malach 2004 The human visual cortex. *Annual Review of Neuroscience* 27: 649-677.

Grossberg, Stephen and Mignolla, Ennio 1985 Neural Dynamics of Form Perception: Boundary Completion, Illusory Figures, and Neon Color Spreading. *Psychological Review 92*: 173-211.

[Hansen, T., Pracejus, L. and Gegenfurtner, K R. Color perception in the intermediate periphery of the visual field.](http://.../%22%20%5Ct%20%22_blank)*[Journal of Vision 9](http://.../%22%20%5Ct%20%22_blank)*[(4):26, (2009), 1–12](http://.../%22%20%5Ct%20%22_blank).

Hardin, C. L. 1988 *Color for Philosophers*. Indianapolis: Hackett.

Hedge’, J. F. Fang, S. O. Murray and D. Kersten 2008 Preferential responses to occluded objects in the human visual cortex. *Journal of Vision* 8: 16-35.

Hertrich, I., Dietrich, S., & Ackermann, H., 2011. Cross-modal interactions during perception of audiovisual speech and nonspeech signals: an fMRI study. *Journal of Cognitive Neuroscience*, 23, 221-237.

Huemer, M. 2001 *Scepticism and the Veil of Perception*. Lanham: Rowman and Littlefield.

Iurilli, G., D. Ghezzi, U. Olcese, G. Lassi, C. Nazzaro, R. Tonini, V. Tucci, F. Benfenati and P. Medini 2012 Sound-driven synaptic inhibition in primary visual cortex. *Neuron* 73: 814-828.

James, T. W., Humphrey, G. K., Gati, J. S., Servos, P. Menon, R. S. & Goodale, M. A. 2002 Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia* 40: 1706-1714.

Katzner, S. and S. Weigelt 2013 Visual cortical networks: Of mice and men. *Current Opinion in Neurobiology* 23: 202-206.

Kentridge, R.W., Heywood, C.A., and Weiskrantz, L. (1999). Attention without awareness in blindsight, *Proc. R. Soc. Lon. B*, 266: 1805-1811

Komatsu H, Kinoshita M, Murakami I. 2000 Neural responses in the retinotopic representation of the blind spot in the macaque V1 to stimuli for perceptual filling-in. *Journal of Neuroscience* 20: 9310-19.

Komatsu, H. 2006 The neural mechanisms of perceptual filling-in. *Nature Review Neuroscience* 7: 220-231.

Kosslyn, S. M., Behrmann, M., & Jeannerod, M. (1995a). The cognitive neuroscience of mental imagery. *Neuropsychologia*, 33, 1335-1344.

Kovacs, G., Vogels, R., & Orban, G. A. (1995). Selectivity of macaque inferior temporal neurons for partially occluded shapes. *Journal of Neuroscience*, 15, 1984-1997.

Lamme, V. A. and P. R. Roelfsema 2000 The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience* 23: 571-579.

Lee, T. S. and Nguyen, M. 2001 Dynamics of subjective contour formation in the early visual cortex. *Proceedings of the National Academy of Sciences 98*: 1907-1911.

Lerner, Y., Harel, M., & Malach, R. (2004). Rapid completion effects in human high-order visual areas. *NeuroImage*, 21, 516–526

# [Li W](http://www.ncbi.nlm.nih.gov/pubmed/?term=Li%20W%5BAuthor%5D&cauthor=true&cauthor_uid=15156149)., [Piëch V](http://www.ncbi.nlm.nih.gov/pubmed/?term=Pi%C3%ABch%20V%5BAuthor%5D&cauthor=true&cauthor_uid=15156149), [Gilbert CD](http://www.ncbi.nlm.nih.gov/pubmed/?term=Gilbert%20CD%5BAuthor%5D&cauthor=true&cauthor_uid=15156149). 2004 Perceptual learning and top-down influences in primary visual cortex. *Nature Neuroscience* 7: 651-7.

Lommertzen, J. ,Van Lier, R., & Meulenbroek, R. (2009). Grasping partly occluded objects: effects of global stimulus information on Action. *Perception*, 38, 200-214.

Lupyan, G., Thompson-Schill, S. L., & Swingley, D. (2010). Conceptual penetration of visual processing. *Psychological Science*, *21*(5), 682–691.

Lupyan, G. (2015). "Cognitive Penetrability of Perception in the Age of Prediction: Predictive Systems are Penetrable Systems." *Review of Philosophy and Psychology*: 1-23.

[Lyons, Jack (2011). Circularity, Reliability, and the Cognitive Penetrability of Perception.](http://philpapers.org/rec/LYOCRA%22%20%5Ct%20%22_blank) *Philosophical Issues*21 (1): 289-311.

[Lyons, Jack C. (2015). Unencapsulated Modules and Perceptual Judgment.](http://philpapers.org/rec/LYOUMA%22%20%5Ct%20%22_blank) In J. Zeimbekis & A. Raftopoulos (eds.), [*The Cognitive Penetrability of Perception: New Philosophical Perspectives*](http://philpapers.org/rec/JZECP). Oxford University Press 103-122.

McGrath, M. 2013 ‘Phenomenal Conservatism and Cognitive Penetration: The Bad Basis Counterexamples.’ In: *Seemings and Justification*. Ed. Chris Tucker. Oxford: Oxford University Press.

 Macpherson, F. 2012 Cognitive penetration of colour experience. *Philosophy and Phenomenological Research* 84: 24-62

Martuzzi, R. Murray, M. M., Michel, C. M., Thiran, J-P., Maeder, P. P., Clarke, S. & Meuli, R. A. 2007 Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cerebral Cortex* 17: 1672-1679.

Murray, S. O., D. Kersten, B. A. Olshausen, P. Schrater and D. L. Woods 2002 Shape perception reduces activity in human primary visual cortex. *PNAS* 99: 15164-15169.

Nanay, Bence 2006 Does what we want influence what we see? *Proceedings of the 28th Annual Conference of the Cognitive Science Society (CogSci 2006)*. Hillsdale, NJ: Lawrence Erlbaum, pp. 615-621.

Nanay, Bence 2009 Shape constancy, not size constancy: a (partial) explanation for the Müller-Lyer illusion. In: N.A. Taatgen & H. van Rijn (eds.): *Proceedings of the 31st Annual Conference of the Cognitive Science Society* *(CogSci 2009)*. Mahwah, NJ: Lawrence Erlbaum, 2009, pp. 579-584.

Nanay, Bence 2010 Perception and Imagination: Amodal Perception as Mental Imagery. *Philosophical Studies* 150: 239-254.

Nanay, Bence 2011 Do we see apples as edible? *Pacific Philosophical Quarterly* 92: 305-322.

Nanay, Bence 2013 *Between Perception and Action*. Oxford: Oxford University Press.

Nanay, Bence 2016 Hallucination as mental imagery. *Journal of Consciousness Studies*. Forthcoming.

Nanay, Bence forthcoming *Seeing Things You Don’t See*. Oxford: Oxford University Press.

Nieder, A. 2002 Seeing more than meets the eye: processing of illusory contours in animals. *Journal of Comparative Physiology A* 188: 294-260.

O’Callaghan, Casey 2008a Object perception: Vision and audition. *Philosophy Compass* 3: 803-829.

O’Callaghan, Casey 2008b "Seeing What You Hear: Cross-Modal Illusions and Perception" *Philosophical Issues* 18(1): 316-338.

O’Callaghan, Casey 2011 Lessons from beyond vision. *Philosophical Studies* 153: 143-160.

O’Callaghan, Casey 2014 Intermodal binding awareness. In: D. Bennett and C. Hill (eds.): *Sensory Integration and the Unity of Consciousness.* Cambridge, MA: MIT Press, pp. 73-96.

O’Connor, D. H., Fukui, M. M., Pinsk, M. A., & Kastner, S. (2002). Attention modulates responses in the human lateral geniculate nucleus. *Nature Neuroscience* 5(11), 1203–1209.

Pearson, J & Westbrook, F. (2015) Phantom perception: voluntary and involuntary nonretinal vision. *Trends in Cognitive Sciences* 19: 278-284.

Pearson, Joel, Thomas Naselaris, Emily A. Holmes, and Stephen M. Kosslyn 2015 Mental Imagery: Functional Mechanisms and Clinical Applications. *Trends in Cognitive Sciences* 19: 590-602.

Pekkola, J., Ojanen, V., Autti, T., Jaaskelainen, I.P., Mottonen, R., Tarkainen, A. & Sams, M. 2005 Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *NeuroReport* 16: 125-128.

Pessoa, L., & Ungerleider, L.G. (2005). Visual attention and emotional perception. In L. Itti, G. Rees, and J.K. Tsotsos (Eds.), *Neurobiology of attention*. San Diego, CA: Elsevier.

Phillips, Ian 2013 Afterimages and sensation. *Philosophy and Phenomenological Research* 87: 417-453.

Pryor, James. 2000. "The Skeptic and the Dogmatist." *Noûs*34(4): 517-49.

Pylyshyn, Zenon 1999 Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. [*Behavioral and Brain Sciences*](http://www.journals.cambridge.org/action/displayAbstract?fromPage=online&aid=30933##) 22: 341-365.

Ramachandran V. S. 1992 Filling in the blind spot. *Nature* 356: 115.

Ramachandran V. S. and Gregory R. 1991 Perceptual filling-in of artificially induced scotomas in human vision. *Nature* 350: 699-702.

Rauschenberger, R. and S. Yantis 2001 Masking unveils pre-amodal completion representation in visual search. *Nature* 410: 369-372.

Rauschenberger R, Liu T, Slotnick SD, Yantis S: Temporally unfolding neural representation of pictorial occlusion. *Psychological Science* 2006,17:358–364.

Rolls, E. T. and M. J. Tovee 1994 Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proceedings of the Royal Society: Biological Sciences* 257: 9-15.

Roseboom, W., T. Kawabe and S. Nishida 2013 The cross-modal double flash illusion depends on featural similarity between cross-modal inducers. *Scientific Reports* 3: 3437-4359.

Scherzer, T. R. & Ekroll, V. 2015 Partial modal completion under occlusion: What do modal and amodal percepts represent. *Journal of Vision* 15: 1-20.

Schmitz, Taylor W., Eve De Rosa, and Adam K. Anderson  2009 Opposing Influences of Affective State Valence on Visual Cortical Encoding *The Journal of Neuroscience* 29: 7199 –7207.

Schupp, Harald T., Bruce N. Cuthbert, Margaret M. Bradley, Charles H. Hillman, Alfons O. Hamm, and Peter J. Lang 2004 Brain processes in emotional perception: Motivated attention. *Cognition and Emotion* 18: 593-611

Sekuler, A. B. & Palmer, S. E. 1992 Perception of partly occluded objects: a microgenetic analysis. *Journal of Experimental Psychology General*  121: 95-111.

Sekuler, R., Sekuler, A. B., and Lau, R. (1997). Sound alters visual motion perception. *Nature*, 285: 308.

Shams, L., Kamitani, Y., and Shimojo, S. 2000 What you see is what you hear. *Nature* 408: 788.

Siegel, S. 2011 Cognitive penetrability and perceptual justification. *Nous* 46: 201-222.

Siegel, S. 2013 ‘The Epistemic Impact of the Etiology of Experience.’ *Philosophical Studies* 162: 697–722.

[Silins, Nicholas 2016. Cognitive Penetration and the Epistemology of Perception.](http://philpapers.org/rec/SILCPA-3%22%20%5Ct%20%22_blank) *Philosophy Compass* 11: 24-42.

Smith, F. W. & Muckli, L. 2010 Nonstimulated early visual areas carry information about surrounding context. *PNAS* 107: 20099-20103.

Sosa, Ernest (1999). How to defeat opposition to Moore, *Philosophical Perspectives* 33: 141–153.

Spence, C. and Driver, J. (2004, eds.): *Crossmodal Space and Crossmodal Attention*, Oxford: Oxford University Press.

Sperandio, I., P. A. Chouinard and M. A. Goodale 2012 Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. *Nature Neuroscience* 15: 540-542.

Spillmann L, Otte T, Hamburger K, Magnussen S. 2006 Perceptual filling-in from the edge of the blind spot. *Vision Research* 46: 4252-7.

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

Sugita, Y. (1999). Grouping of image fragments in primary visual cortex. *Nature*, 401, 269-272.

Talavage, T. M., Sereno, M. I..; Melcher, J. R.; Ledden, P.J.; Rosen, B.R.; Dale, A.M. 2004. ["Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity."](http://www.cogsci.ucsd.edu/~sereno/papers/AuditoryMaps04.pdf)  *Journal of Neurophysiology*91 (3): 1282–96.

Thorpe, S. D. Fize and C. Marlot 1996 Speed of processing in the human visual system. Nature 381: 520-522.

Troncoso XG, Macknik SL, Martinez-Conde S. 2008 Microsaccades counteract perceptual filling-in. *Journal of Vision* 8: 15-19.

Tucker, C. 2014 ‘If Dogmatists Have a Problem with Cognitive Penetration, You Do Too.’ *Dialectica* 68: 35–62

Van Essen, D.C. (2004) Organization of visual areas in macaque and human cerebral cortex. In: *The Visual Neurosciences*. L. Chalupa and J.S. Werner, eds., MIT Press, pp. 507-521.

Vance, J. 2015 ‘Cognitive Penetration and the Tribunal of Experience.’ *Review of Philosophy and Psychology* 7: 1–23.

Vetter P., Smith F. W., Muckli L. (2014). Decoding sound and imagery content in early visual cortex. *Current Biology* 24: 1256–1262.

Vrins, S., De Wit, T.C.J., & Van Lier, R. (2009). Bricks, Butter, and Slices of Cucumber: Investigating semantic influences in amodal completion. *Perception*, 38, 17-29.

Vroomen, J., Bertelson, P., and de Gelder, B. (2001) Auditory-visual spatial interactions: Automatic versus intentional components. In: de Gelder, B., de Haan, E., and Heywood, C. (eds): *Out of Mind*, Oxford: Oxford University Press, pp. 140–150.

Vul, E. et al. (2008) The McCollough effect reflects permanent and transient adaptation in early visual cortex. *Journal of Vision* 8, 4

Watanabe, T. and Sato, T. (1989) Effects of luminance contrast on color spreading and illusory contour in the neon color spreading effect. *Perception and Psychophysics* 45, 427–430

Watkins, S., Shams, L., Tanaka, S., Haynes, J. D., and Rees, G. 2006 Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage* 31: 1247-1256.

Weil RS, Kilner JM, Haynes JD, Rees G. 2007 Neural correlates of perceptual filling-in of an artificial scotoma in humans. *Proceedings of the National Academy of Sciences U S A* 104: 5211-16.

Weil RS, Watkins S, Rees G. 2008 Neural correlates of perceptual completion of an artificial scotoma in human visual cortex measured using functional MRI. *Neuroimage* 42: 1519-28.

Welchman AE, Harris JM. 2001 Filling-in the details on perceptual fading. Vision Research 41: 2107-17.

Zangaladze, A., Weisser, V. D., Stilla, E., Prather, S. C. & Sathian, K. 1999 Involvement of visual cortex in tactile discrimination of orientation. *Nature* 401: 587-590.

1. Roberto Bolano: *2666*. London: Picador, 2009, p. 219. [↑](#footnote-ref-1)
2. I should note that some dogmatists take a different route to argue that the findings about cognitive penetration do not threaten their account of perceptual justification (see esp. Pryor 2000, Huemer 2013, see also Huemer 2006). The focus of this paper is not dogmatism. [↑](#footnote-ref-2)
3. One example of this is the (very complicated) correspondence between the retina and V4/V8, when it comes to color, see Hardin 1988 for a philosophical overview. [↑](#footnote-ref-3)
4. One may object: hasn’t Daniel Dennett’s repeated skepticism about ‘filling-in’ the blind spot (Dennett 1991, p. 335ff) demonstrated that this story is incorrect? I don’t think so. First, there is plenty of empirical evidence that the early cortices do actively ‘fill-in’ the missing part of the visual scene (see, for example, Komatsu et al. 2000 and also Churchland and Ramachandran 1993 – and also Akins and Winger 1996 for a very good overview of this debate). Second, I’m not even sure that Dennett would disagree with anything I say here – his concern in Dennett 1991 was about phenomenology – whether there is conscious filling in. And I’m certainly not arguing that there is – there is cortical filling in. The mental imagery involved in the filling in of the blind spot is almost always unconscious. [↑](#footnote-ref-4)
5. Note that the term ‘amodal’ is a bit of a misnomer here: the completion by any account happens visually – the term ‘amodal’, traditionally, was supposed to indicate that this process is not triggered by sensory stimulation. [↑](#footnote-ref-5)
6. Much of this section is about the intricate connections between different sense modalities. Nonetheless, in the definition of mental imagery in general and of multimodal mental imagery in particular, I am relying on the difference between perceptual processing in different sense modalities. It is important to emphasize that there is no tension between these two claims – in spite of all the intricate links between the perceptual processing in different sense modalities, we can nonetheless identify what distinctively visual perceptual processing amounts to. Multimodality does not imply that there are no distinct sense modalities. [↑](#footnote-ref-6)
7. This argument clearly does not apply to dogmatist theories of perceptual justification – I am not discussing dogmatism about perceptual justification in this paper. [↑](#footnote-ref-7)