Cognitive Penetration: Inference or Fabrication?

Lu Teng, New York University Shanghai, lt77@nyu.edu

(This is a penultimate draft. Please cite the published version)

Abstract: Cognitive penetrability refers to the possibility that perceptual experiences are influenced by our beliefs, expectations, emotions, or other personal-level mental states. In this paper, I focus on the epistemological implication of cognitive penetration and examine how exactly aetiologies matter to the justificatory power of perceptual experiences. I examine a prominent theory, according to which some cognitively penetrated perceptual experiences are like conclusions of bad inferences. Whereas one version of this theory is psychologically implausible, the other version has sceptical consequences. In the second half of the paper, I suggest an alternative theory, drawing on recent empirical research on imagining-perception interaction and the epistemology of imagining.

Keywords: cognitive penetration; perception; perceptual inference; sensory imagining; justification; epistemology

1. Introduction

In psychology and the philosophy of mind, an important problem about perception concerns its cognitive penetrability—roughly speaking, whether perceptual experiences can be influenced by our beliefs, expectations, emotions, or other personal-level mental states, when the following factors are fixed: the distal stimuli, external conditions such as lighting, conditions of our sensory organs, and focus and spatial attention. In Hansen et al. [2006],

subjects were shown a picture of a yellow banana, and were asked to adjust its colour until it looked achromatic/entirely grey. As it turned out, the subjects adjusted the colour slightly in the direction of blue, the opposite direction from yellow, which suggested that at the point where the picture was entirely grey, the subjects still saw it as a bit more yellow. This effect, according to one plausible explanation, might be due to the influence of the subjects' background cognition that bananas are typically yellow.

The problem about the cognitive penetrability of perception has also drawn a lot of interest from epistemologists: what is the epistemological implication if cognitive penetration ever happens? If cognitive penetration happened in Hansen et al. [2006]'s experiment above, did the subjects' experience give them the same amount of justification for believing that the banana picture was yellowish-grey—when it was entirely grey—as a cognitively unpenetrated yellowish-grey experience would normally give? According to the Downgrade Thesis, some cognitively penetrated perceptual experiences are epistemically downgraded, in the sense that they give the subjects less justification for believing their penetrated contents than perceptual experiences that are unpenetrated to represent the same contents would normally give.

Many epistemologists subscribe to the Downgrade Thesis.[1] But this conclusion needs explanation: what makes some cognitively penetrated experiences epistemically downgraded? In this paper, I assume that perceptual experiences are cognitively penetrable so that we can focus on the epistemological matter. I first examine a prominent explanation in the literature, according to which some cognitively penetrated perceptual experiences are like conclusions of bad inferences: just as beliefs resulting from bad inferences are less justified, perceptual experiences resulting from bad inferences also have less justificatory power. I call this theory

[1] For some defenses of the Downgrade Thesis, see Ghijsen [2016], Long [2018], Lyons [2011, 2016], McGrath [2013], Munton [2019], Siegel [2012, 2013, 2017], Teng [2016], and Vance [2014].

inferentialism.[2] In spelling out the theory, inferentialists focus on different kinds of perceptual inference. I argue that all of the existing versions of inferentialism face serious problems.

My alternative explanation draws inspiration from the epistemology of imagining, where 'imagining' generally refers to perception-like experiences that are not triggered by the relevant external stimuli, or mental processes that generate these experiences. Following Macpherson [2012], I argue that it is psychologically plausible for cognitive penetration to happen through imagining-perception interaction. Some imaginings lack justificatory power. This is importantly because they are experiences we *fabricate*. I apply this theory to cognitive penetration, and suggest that, given the role imaginings play in cognitive penetration, some cognitively penetrated perceptual experiences are also epistemically downgraded. I call this alternative theory *fabricationism*.

In sections 2 to 3, I introduce two versions of inferentialism. I argue that whereas one version is psychologically implausible, the other version may lead to unacceptable scepticism about perceptual justification. In sections 4 and 5, I propose and defend fabricationism.

2. Experiential Inferentialism

Consider two hypothetical cognitive penetration cases:

> **Anger:** Jill fears that Jack is angry with her. When she meets Jack, her fear causes her to see Jack's face as expressing anger, although Jack is not angry at all.
>
> **Preformationism:** Preformationists irrationally believed that sperm cells contain embryos. When they observed sperm cells under a microscope, their background belief caused them to see embryos in the cells.[3]

[2] Some adherents of inferentialism include McGrath [2013] and Siegel [2013, 2017].

The perceptual experiences in these cases seem epistemically downgraded. Jill's visual experience seems to give her less justification for believing that Jack has an angry face than a cognitively unpenetrated anger experience would normally give. The preformationists' visual experience seemed to give them less justification for believing that the observed sperm cells contained embryos than a cognitively unpenetrated embryo experience would normally give. To explain these conclusions, inferentialists propose two theses:

> **The Descriptive Thesis:** Perceptual experiences can result from inferences.
>
> **The Epistemic Thesis:** The justificatory power of perceptual experiences can be influenced by the quality of the inferences leading to them.

Inferentialists explain that, in the two cases above, the perceptual experiences can be seen as conclusions of bad inferences. That is why they are epistemically downgraded.

To spell out this idea, different inferentialists treat different transitions as perceptual inferences. McGrath [2013] models perceptual inferences on belief inferences. He points out that, in order for a transition to be an inference, the inputs and outputs must be mental states attributed to the subject, and the transition must be rationalised by the subject's own mental states. Based on this characterization, McGrath claims that a transition from one experience to another experience is 'quasi-inferential' just in case, had we replaced the experiences with corresponding beliefs, the resulting transition would be an inference by the subject.

McGrath argues that some cognitive penetration cases involve quasi-inferences from one experience to another experience. For example, in the anger case, Jill's fear might cause her to infer from an experience that Jack's facial muscles group in a certain way (which does not express anger) to another experience that Jack has an angry face. In the preformationism

[3] These two cases are from Siegel [2012].

case, the preformationists' background belief that sperm cells contain embryos might cause them to infer from an experience that the observed sperm cells had components of a certain shape, size, and colour (which were not good signs of embryos) to another experience that the cells contained embryos.

McGrath points out that there are at least two different ways for a belief inference to be bad. First, if the premise is unjustified, then the conclusion inherits the unjustifiedness. Second, if the premise does not support the conclusion well, then the inference jumps to the conclusion and the conclusion is unjustified. McGrath thinks that perceptual inferences are also susceptible to these kinds of epistemic inappropriateness. For example, Jill's fear fails to bridge the gap between the input and output experiences. This perceptual inference jumps to the conclusion, and so Jill's output experience is epistemically downgraded. The preformationists' background belief was unjustified. Their output experience inherited the unjustifiedness, and hence was also epistemically downgraded.

One objection to McGrath's view is that he focuses on transitions from an experience that represents low-level properties, such as shape, size, and colour, to another experience that represents high-level properties, such as being angry, or being an embryo. However, there could be cognitive penetration to experiences of low-level properties. Consider a simplified case based on Hansen et al. [2006]'s experiment (as explained earlier, the subjects, when asked to adjust a yellow banana picture until it looked entirely grey, adjusted its colour slightly in the direction of blue. This meant that at the point when the picture was entirely grey, the subjects still saw it as yellowish-grey):

> **Banana:** A banana picture was entirely grey, but the subjects' background belief that banana-shaped objects are typically yellow caused them to see it as yellowish-grey. The subjects' experience seemed to give them less justification for believing

that the picture was yellowish-grey than a cognitively unpenetrated yellowish-grey experience would normally give.

McGrath admits this challenge. He replies [2013:239n25] that the subjects' background belief might cause them to transition from an experience that the picture was entirely grey to another experience that it was yellowish-grey. If there were no other intermediate steps, then this perceptual inference jumped to the conclusion.

My objection to McGrath's view questions the psychological plausibility of postulating that perceptual inferences are between two experiences. First, consider his explanation of the anger case and preformationism case. The explanation seems to imply that we have an experience of high-level properties, following an experience of low-level properties. But this is probably false. When we see someone as angry in everyday settings, we do not seem to go through two successive experiences, of which the first represents the person's facial muscles grouping in a certain way, and the second experience represents the face as expressing anger. We simply experience an angry face.

McGrath might reply that the input and output experiences can be sub-experiences of the same experience and occur simultaneously, but they nonetheless bear some inference-like relationship with each other (see [ibid.: 237n21]). Accordingly, he does not have to assume that we experience high-level properties after experiencing low-level properties; they can be different aspects of one experience. I acknowledge that the objection above is not devastating to McGrath's account of the anger case and preformationism case. However, there is a more serious concern about the psychological plausibility of his account of the banana case.

In the banana case, the input and output experiences represented incompatible colour properties—being entirely grey and yellowish grey, respectively. These experiences could not

co-exist in one experience. Therefore, there must be a transition between them if McGrath's account is correct.

Now given the phenomenal discrepancy between pure grey and yellowish-grey, it is reasonable to expect the subjects to notice the transition. However, there has been no good experimental evidence. In Hansen et al. [2006], the subjects did not make any verbal reports. Had they successively gone through a grey experience and a yellowish-grey experience at the point when the banana picture was entirely grey, they would have first stopped adjusting the picture in responding to the grey experience, and then would have restarted the adjustment process in responding to the yellowish-grey experience. But Hansen et al. [2006] and related studies [Delk and Fillenbaum 1965; Olkkonen at al. 2008] have not reported observing behaviour patterns like this.[4]

In another widely discussed study, Levin and Banaji [2006] asked subjects to adjust the luminance of a grey patch to match a racially ambiguous face that was labelled either 'Black' or 'White'. When the face was labelled 'Black', the subjects consistently adjusted the patch to a darker shade of grey than the face's actual colour. This effect has been taken to show that the subjects' cognition about racial stereotypes led them to view the 'Black' labelled face as darker. Notably, Levin and Banaji [2006] pointed out that the adjustable patch was originally displayed in a different shade of grey from the 'Black' labelled face, and that sometimes the subjects did not adjust the patch at all before settling on a final decision. This also supported the diagnosis that they simply saw the 'Black' labelled face in a darker grey, rather than moving through two colour experiences. Otherwise, they wouldn't have made zero adjustment.

[4] While I developed these objections on my own, similar discussions can be found in Long [2018].

Our discussion shows that McGrath's version of inferentialism at least needs revision. Even if it is psychologically plausible to postulate that perceptual inferences are between two experiences in the anger case and preformationism case, it is not plausible with respect to the banana case.

3. Subpersonal Inferentialism

Siegel [2017] takes there to be different kinds of perceptual inference, among which inference between two experiences is one kind. In particular, she characterizes inference as 'a distinctive kind of response to an informational state, or to a combination of such states, that produces a conclusion', where the distinctive kind of response need not be conscious, or involve a 'reckoning state' about the support relationship between the input and output states [ibid. 77]. This characterization does not preclude perceptual experiences from resulting from inferences.

Siegel thinks that in order for a kind of mental state to be eligible for being inferential inputs, it must be capable of having justificatory power, or be epistemically evaluable. In the book, she argues that some perceptual experiences are 'epistemically charged', which allows them to have justificatory power and to manifest an epistemic status, and so perceptual experiences are not precluded from being inferential inputs.[5] What makes her view distinctive is that she further argues that unconscious representational states posited by psychological explanations of perception are capable of having justificatory power and epistemically evaluable, and so these states are also not precluded from being inferential inputs.[6]

[5] For the relevant arguments, see Siegel [2017: sec. 3.1].
[6] For the relevant arguments, see Siegel [2017: sec. 5.6].

Siegel explains the epistemic downgrade in the anger case and preformationism case in a similar way to how McGrath does. With respect to the banana case, she does not take it for granted that the subjects' output experience that the banana picture was yellowish-grey was epistemically downgraded. She thinks that this case could be spelled out in multiple ways and its epistemological implication depends on which specific perceptual inference was involved. Suppose that, in this case, the output experience was inferred from an unconscious perceptual state that the banana picture was entirely grey and the subjects' background belief that banana-shaped objects are typically yellow. Moreover, suppose that there were no other intermediate steps. Then this perceptual inference jumped to the conclusion, and the output experience was epistemically downgraded.[7]

By allowing different kinds of mental state to be eligible for being inferential inputs, Siegel's version of inferentialism explains more cognitive penetration cases than McGrath's. Her account of the banana case does not postulate two incompatible colour experiences, and avoids the problem that faces McGrath's account. This is a great advantage of Siegel's view. However, the perceptual system might be limited in various ways. By allowing subpersonal inferences to influence the justificatory power of perceptual experiences, her view might end up implying that numerous perceptual experiences are epistemically downgraded, and hence have sceptical consequences for perceptual justification. In the rest of this section, I develop this worry from a flourishing empirical project, which analyses perception in Bayesian terms. I choose Bayesian theories of perception partly because Siegel discusses them. But the worry needs not to hinge on their truth.

The perceptual system is known to face an underdetermination problem—the sensory data received by the perceptual system are insufficient to determine their external causes. For

[7] For a detailed examination of the banana case, see Siegel [2017: sec. 7.2].

example, the sensory data received when viewing a convex object are consistent with at least two possibilities: (i) the object is convex and illuminated from above, or (ii) it is concave and illuminated from below. However, the perceptual system has to figure out what experiences to generate. According to Bayesian theories of perception, the perceptual system solves this problem through a two-step inference. In the first step, for all of the relevant hypotheses about the cause of the sensory data, the perceptual system calculates their posterior probabilities based on stored assumptions about the environment and by following Bayes' Theorem.

The orthodox theory of perceptual experience treats the latter as non-probabilistic: we perceive things as having a certain shape, size, colour, and so forth rather than a spectrum of these properties. In the second step, the perceptual system selects a hypothesis among all of the relevant hypotheses to generate a non-probabilistic perceptual experience. The selection is based on probabilities from the previous calculation and follows a certain decision rule, such as 'Select the hypothesis with the highest posterior probability'. In the convexity-concavity case, with a stored assumption that lighting usually comes from above, the perceptual system assigns a higher posterior probability to 'The object is convex and illuminated from above' than the alternative hypothesis. The perceptual system then generates an experience that the object is convex and illuminated from above.

If Bayesian theories of perception are true, then Siegel's inferentialism can be applied to a Bayesian perceptual inference. I focus on the second step of the inference. Suppose that the selected hypothesis is H and its posterior probability is *n*. This step seems to move from an unconscious perceptual state that contains probabilistic information about H to a non-probabilistic experience that H. Even if the posterior probability of H, *n*, is quite high, by representing H simply as true in the output experience, this inference fails to take account of

the probabilistic nuance in the premise. It seems to jump to the conclusion. Consider that, in the convexity-concavity case, the output experience seems to fully commit to 'The object is convex and illuminated from above'. The experience embodies overconfidence that is not properly supported by the premise; the latter treats the hypothesis as its best guess, without full commitment.

To further buttress this objection, consider an example that involves a corresponding belief inference. In a murder, the police have tracked down five suspects. Given the obtained evidence, suspect A probably did it, but the posterior probability of this hypothesis is only 30%. It is epistemically inappropriate for the police to infer that suspect A committed the murder. The move obviously jumps to the conclusion. Even if the posterior probability of the hypothesis is much higher—say, 95%, forming a full belief that A committed the murder seems epistemically inappropriate. There still seems to be overconfidence embedded in the full belief. If inferences from beliefs about probabilities to full beliefs typically jump to the conclusion, then isomorphic inferences from unconscious perceptual states about probabilities to non-probabilistic experiences also typically jump to the conclusion.

Bayesian theories of perception combined with Siegel's version of inferentialism seem to imply that a large number of perceptual experiences result from bad inferences, and hence are epistemically downgraded. This may lead to unacceptable scepticism about perceptual justification, which gives us a good reason to reject Siegel's view.

Before turning to my proposal, I consider three replies. First, one might acknowledge that the perceptual system has limitations, and suggest that perceptual inferences be subject to different standards of evaluation from belief inferences. My main worry about this reply is that inferentialism would lose its initial appeal, where the latter partly consists in the fact that we have a relatively clear idea about how to evaluate belief inferences. If perceptual

inferences were subject to different standards, then inferentialists would need to explain what the standards are, and why these are plausible.

Second, our discussion assumes that perceptual experiences are non-probabilistic, but this assumption might be rejected. Morrison [2016] and Munton [2016] argue that perceptual experiences are probabilistic. One reason they provide is that this theory better explains why fully trusting our perceptual experiences leads to different credences. One might suggest that if perceptual experiences inherit their probabilistic component from the Bayesian calculation, then the second step of a Bayesian perceptual inference can be epistemically appropriate.

The probabilistic approach to perceptual experience is controversial.[8] Even if its truth is granted, then, unless perceptual experiences inherit the probabilistic component loyally from the Bayesian calculation, a similar worry arises. After all, a move from a state that H's posterior probability is *n* to a perceptual experience that assigns a different probability to H might still be problematic. Suppose that the experience does assign a probability of *n* to H. The second step of a Bayesian perceptual inference then seems fine. But, for many Bayesian theorists, the calculation of posterior probabilities at best follows Bayes' Theorem approximately. Siegel's version of inferentialism might still predict that this step is bad and the output experience is epistemically downgraded. The probabilistic approach is harder to pursue than expected.

It is worth stressing again that the sceptical worry about Siegel's view need not hinge on the truth of Bayesian theories, which already analyses perception in an idealised way. It is entirely possible that the perceptual system engages in less than optimal mental processes. It might follow various heuristics, which are useful but fall short of standards required by logic

[8] For a critical discussion, see Siegel [ms].

and rationality. When applied to such perceptual processes, Siegel's view might more straightforwardly lead to extensive scepticism.

Finally, one might admit that Siegel goes too far by allowing subpersonal-level mental states to be inferential inputs. Whereas McGrath is correct in thinking that inferential inputs and outputs should be mental states of the subject, his theory is too narrow by focusing on inferences between experiences. One might therefore propose a middle ground: namely, we allow personal-level mental states other than experiences and beliefs to be inferential inputs, but we preclude subpersonal-level mental states.

This approach encounters problems in explaining cases where cognitive penetration happens to experiences that represent low-level properties, while the penetrating states are mental states like emotions and desires. As an illustration, suppose that your desire that there is a yellow banana makes you see a banana picture as yellowish-grey when it is in fact entirely grey.[9] In order for inferentialism to apply to this case, it should involve a genuine inference. The middle-ground approach might claim that your experience is inferred from your desire.[10] But it is not clear whether desires can be inferential inputs. We explained that Siegel requires inferential inputs to have justificatory power or to be epistemically evaluable. Even if desires are epistemically evaluable, there might be other necessary conditions, such as representing their contents as true, that they fail to satisfy. One important consideration is that, in an inference, the premise is treated as a reason for the conclusion. If a mental state does not represent its content as true, then it is not clear that the state can function as a reason on the basis of which the conclusion is drawn.[11]

[9] For a recent philosophical defense of the penetrability of desires to perceptual experiences, see Stokes [2012].
[10] For the reasons discussed in section 2, the perceptual inference cannot be between experiences that represent incompatible low-level properties.
[11] Also, consider that beliefs caused by desires are said to be the result of wishful thinking, not inference.

4. Cognitive Penetration Through Imagining

Inferentialism takes cognitive penetration to involve perceptual inferences. However, there might be alternative psychological mechanisms. In this section, I introduce one such mechanism. In the next section, I provide an explanation of the epistemological implication of cognitive penetration in light of this mechanism.

Macpherson [2012] draws our attention to the psychological facts that our personal-level mental states such as beliefs, expectations, and emotions can cause sensory imaginings, and that sensory imaginings can interact with perceptions to generate experiences that have content and phenomenal contribution from both sources. The former psychological fact is widely recognised. To give an example, your expectation that it will snow can cause you to imagine seeing snow when looking out the window, and this can occur either deliberately or spontaneously.

To support the possibility of imagining-perception interaction, Macpherson appeals to a sequence of experiments conducted by Perky [1910] and Segal [1971, 1972], in which subjects were requested to imagine various objects while being secretly presented with faintly coloured pictures. Many subjects reported experiences that seemed to have direct content and phenomenal contribution from both their imaginings and their perceptions of the presented pictures. In one experiment, the subjects were requested to imagine a city skyline while being secretly shown a faint red tomato. Several subjects reported having an imagery of Manhattan at Sunset, which seemed to result from an interaction between the subjects' imagining of the skyline of Manhattan and their perception of the faint red tomato [Segal 1972].

Based on these psychological facts, Macpherson proposes a mechanism of cognitive penetration: in some cognitive penetration, one's personal-level mental states such as beliefs,

expectations, and emotions start an imaginative process; this interacts with one's perceptual process to generate a perceptual experience that has content and phenomenal contribution from both processes. According to this mechanism, what happened in the banana case from Hansen et al. [2006] might be: at the point when the banana picture was in fact entirely grey, the subjects' belief that bananas are typically yellow activated, from top down, an imaginative process that would give rise to an experience that the banana picture was yellow; this process interacted with a bottom-up perceptual process that would give rise to an experience that the banana picture was entirely grey, resulting in the subjects' experiencing the banana picture as yellowish-grey.

This mechanism also makes the anger case and preformationism case psychologically explainable. In the anger case, Jack's face is not angry, but Jill's fear triggers an imaginative process that would give rise to an experience that Jack has an angry face. This interacts with Jill's perceptual process, and contributes an anger component to her perceptual experience. In the preformationism case, the observed sperm cells did not contain any embryo, but the preformationists' unjustified background belief triggered an imaginative process that would give rise to an experience that the observed sperm cells contained embryos. This imaginative process interacted with the preformationists' perceptual process, and contributed an embryo component to their resulting perceptual experience.

More clarifications can be made about this mechanism. First of all, roughly speaking, sensory imagining refers to perception-like experiences that are not triggered by the relevant external stimuli, or mental processes that generate these experiences. Given that subjects of cognitive penetration are unaware of enjoying two separate phenomenal states, the kind of imagining-perception interaction involved in cognitive penetration is probably between the imaginative and perceptual processes, rather than two experiences.

Sensory imaginings can occur either deliberately or spontaneously. Some examples of the latter include daydreaming, flashbacks to past scenes, and earworms (in which music sounds pop repeatedly into one's mind). When personal-level mental states such as beliefs, expectations, and emotions penetrate one's perception through imagining, the former states typically activate the imagining spontaneously. Macpherson's mechanism is not committed to the view that just entertaining some cognitive states is sufficient for cognitive penetration to happen, or that we can wilfully manipulate the contents and phenomenal character of our perceptual experiences through imagining. Specific conditions are needed in order for the relevant kind of imagining-perception interaction to take place, which are undetermined at the moment.

Second, in addition to the experiments conducted by Perky and Segal, there is further evidence from recent empirical research that supports the possibility of imagining-perception interaction. One set of evidence comes from the research on multisensory integration. In the cross-bounce illusion, a classical illustration of multisensory integration, subjects watched an ambiguous presentation, in which two beads moved toward each other. This could either be viewed as the beads going across each other, or as their bouncing off each other. Interestingly, if the subjects heard a 'click' sound at the time when the beads encountered each other, then they tended to see the beads as bouncing off each other. This seemed to be due to an interaction between their visual and auditory modalities. Recently, Berger and Ehrsson [2013, 2014] demonstrated that merely imagining hearing a 'click' sound led to a similar effect.

In philosophy, Nanay [2018, ms] argues that multisensory integration cases like the original cross-bounce illusion already involve imagining-perception interaction. In particular, Nanay thinks that in the original cross-bounce illusion, hearing the 'click' sound altered the

subjects' visual experience through cross-modally triggering a visual imagining that the beads bounced off each other. He names this kind of imagining 'multimodal mental imagery', and argues that it is an important component of our everyday perceptions. Nanay's discussion of multimodal mental imageries is consistent with, and further buttresses, the point that sensory imagining can interact with perception.[12]

Another set of evidence is from neuroscience. It has been found that visual imagining and perception activate common areas in the visual cortex. Stokes et al. [2009] revealed that visually imagining X and O activated different areas in the lateral occipital cortex (important for shape discernment), and these areas roughly correspond to the areas activated in visually perceiving X and O, respectively. Consistent with these findings, Reddy et al. [2010] showed that visually imagining and perceiving objects from four different categories—face, building, food, and tool—involved, respectively, common areas in the ventral temporal cortex (critical for visual categorisation). Besides, although it is still controversial, several studies found that visual imaginings even activate the early visual cortex [Kosslyn and Thompson 2003; Page et al. 2011; Pearson and Kosslyn 2015].

These neuroscientific findings are significant. If our neural processes physically realise our mental processes, then the findings support that sensory imagining and perception have common access to highly specific brain areas, and hence there is a neural basis for imagining-perception interaction. Theoretically speaking, the interaction can take different forms. First, as in the Perky effect, the perceptual process might activate brain areas that are not activated by the imaginative process, and directly contribute to the contents and phenomenal character of the subject's imaginative experience. Contrariwise, the imaginative process might activate certain brain areas that are not activated by the perceptual process and

[12] For more philosophical discussions of the cross-modal effects of sensory imaginings, see Spence and Deroy [2013] and Stokes and Biggs [2014].

influence the subject's perceptual experience. This might be the underlying mechanism of the imagining-perception interaction Macpherson has in mind. Finally, the two processes might both activate the same brain areas and have priming effects on each other, to which I come back in a little bit.

The third clarification I want to make is that Macpherson's mechanism has room for apparently innocent or good cognitive penetration. Consider the following two cases:

> **Friend:** You are waiting for a best friend at the train station. Your desire to see this friend makes you spot her more quickly in the crowd. Your visual experience seems to give you justification for believing that the person is your friend.
>
> **Birdwatcher:** When an expert birdwatcher views an eastern kingbird, her expertise makes her see it as a kingbird, whereas a novice might not recognise it. The expert's visual experience seems to give her justification for believing that it is a kingbird.[13]

Macpherson focuses on a particular kind of imagining-perception interaction: namely, the imaginative and perceptual processes directly contribute to the contents and phenomenal character of the subject's perceptual experiences. However, as mentioned above, there might be other forms of imagining-perception interaction, which could be coherently incorporated into Macpherson's mechanism. Pearson et al. [2008] and Pearson et al. [2011] recently found that sensory imagining could have facilitatory effects on perception. Their experiments were about binocular rivalry—the phenomenon that, when one's left eye and right eye are presented with two different pictures (for example, the left eye is presented with a green vertical grating, and the right eye is presented with a red horizontal grating), one is only aware of one picture (such as the green vertical grating) at a time. Pearson and colleagues

[13] For a recent philosophical defense of perceptual expertise involving cognitive penetration, see Stokes [forthcoming].

showed that whereas, under passive viewing, subjects tended to become aware of the picture that was dominant in the previous trial, imagining one of the pictures during a blank intervening period made the subjects more disposed to become aware of the same picture in the subsequent trial, regardless of whether the picture was previously dominant or suppressed.

When considering the underlying neural basis for imagining-perception interaction, I argued that since imaginative and perceptual processes activate overlapping brain areas, it is possible for these processes to prime each other. Pearson et al. [2008: 986] suggest a similar point in discussing the facilitatory effects of imagining on the perception of binocular rivalry: 'It is striking to consider that "perceptual priming" might be elicited without perception, by simply imagining a previous perceptual experience.' If the friend case and birdwatcher case involve cognitive penetration, then one possible mechanism is: the subjects' beliefs, desires, expectations, or other personal-level mental states cause an imaginative process, which then primes their perceptual system to be more sensitive to certain external stimuli. This kind of imagining-perception interaction could be incorporated into Macpherson's mechanism.[14]

5. Fabrication in Imagining and Cognitive Penetration

In this section, I suggest an alternative explanation of the epistemological implication of cognitive penetration. First, I argue that some sensory imaginings lack justificatory power importantly because they are experiences we fabricate. To be consistent, the kind of belief I focus on is non-modal (such as 'It is snowing') rather than modal (such as 'It can be snowing'). Then I apply this theory to cognitive penetration.

[14] In addition to imagining-perception interaction, there might be other mechanisms of apparently innocent or good cognitive penetration. For some recent philosophical discussions of cognitive penetration through object-based and feature-based attention, see Mole [2015], Stokes [2018], and Wu [2017].

It is uncontroversial that some sensory imaginings lack justificatory power. What is controversial is why they lack it. One thought might be that only mental states with a certain attitude—namely representing their contents as true—can have justificatory power. Sensory imaginings represent their contents as merely possible [Hart 1988; Yablo 1993], or they take no attitude at all toward their contents [Kriegel 2015; McGinn 2004; Searle 1983].

This theory overlooks the representational power of sensory imaginings. As Langland-Hassan [2015] convincingly argues, at least some sensory imaginings seem to represent their contents as true. One example is that you saw many misleading pictures of Arc de Triomphe in the past, which represents its colour as silver, and, before going to Paris, you imagine seeing a big silver arc. Your imagining seems to represent as true that Arc de Triomphe is silver. To see this point, consider that, when you visit Arc de Triomphe in person and see its real colour, you will consider your imagining as a mis-imagining. However, if your imagining represented as merely possible that Arc de Triomphe is silver, or entertained the content without taking any attitude, then your imagining would not be a mis-imagining.

Kind [2016, 2018] argues that there can be epistemically good sensory imaginings. Consider that you try to figure out whether an Ikea sofa fits into a space in your living room by imagining moving the former to the latter. If your imagining shows that the sofa fits into the space, then it seems to give you some justification for believing so. Or consider that you try to decide what a picture looks like when rotated 90-degrees clockwise by mentally rotating the picture 90-degrees clockwise. If your imagining suggests that the rotated picture looks like *this*, then your imagining also seems to give you some justification for believing so. If mental states need to represent their contents as true in order to have justificatory power, then there is a further epistemological argument that sensory imaginings can represent their contents as true.

Moreover, if cognitive penetration happens through imagining-perception interaction, where the imaginative component is confused for ordinary perception, then the imaginative component is likely to be represented as true.[15] Once we realise that sensory imaginings can represent their contents as true, we focus on such cases and examine why some of them lack justificatory power. Let's consider the following case:

> **Blizzard:** Bob broke up with his girlfriend, and does not want to go to class. In the past, Bob's school always closed during severe weather, so he hopes that a blizzard will arrive soon. When he looks through the window this morning, this hope causes him to imagine seeing snow. Obviously, Bob's imagining does not give him justification for believing that it is snowing.

In the rest of this paper, I assume that Bob's imagining represents as true that it is snowing. I also assume that Bob's hope can either deliberately or spontaneously cause his imagining.

My preliminary proposal is that Bob's imagining lacks justificatory power importantly because it lacks a good evidential basis: he hopes that a blizzard will reach the town soon, but he has no evidence for it. This proposal, however, needs qualification. Many experiences are not generated on a good evidential basis, but they seem to have justificatory power. One example is perceptual experiences that are not cognitively penetrated. The perceptual system gives rise to them quite automatically. Moreover, in the Müller-Lyer illusion, even if we possess prior evidence that the two lines are of the same length, the perceptual system generates this illusion in us anyway. Since ordinary perceptions and illusions are taken to give us defeasible justification, the requirement for a good evidential basis should not apply to them.

[15] For more examples of sensory imaginings confused for ordinary perceptions, see Teng [2018].

In light of these considerations, I propose to modify the preliminary proposal, to the following: when a mental state results from a personal-level mental process, the mental state needs a good evidential basis in order to have justificatory power; on the other hand, when a mental state results from a merely subpersonal-level mental process, it does not need a good evidential basis in order to have justificatory power. An experience is one we *fabricate* when it results from a personal-level mental process but lacks a good evidential basis. In the blizzard case, Bob's imagining is a fabricated experience, and hence fails to give Bob justification. To support this proposal, I next explain what the personal-subpersonal distinction is, and why it matters to justification.

The personal-subpersonal distinction is not an easy distinction to pin down, although it is usually taken for granted. As originally introduced by Dennett [1969], this distinction is one of psychological explanations. Personal-level psychological explanations ascribe mental states and processes to the whole person. For example, a folk-psychological explanation that posits beliefs and desires and their interactions is a representative personal-level explanation. Subpersonal-level psychological explanations employ functional analysis to decompose the whole person into numerous functional components or subsystems, and ascribe mental states and processes to the latter. For example, a computational theory of depth perception that posits a series of unconscious representations and transitions internal to the perceptual system is a typical subpersonal-level explanation.

Following this distinction of psychological explanations, I take personal-level mental states and processes to refer to states and processes posited by personal-level psychological explanations, and subpersonal-level mental states and processes to refer to those posited by subpersonal-level psychological explanations. Drayson [2012, 2014] points out that, under certain hypotheses about the relationship between personal- and subpersonal-level

psychological explanations, some, but not all, subpersonal-level mental states and processes correspond to personal-level mental states and processes [Fodor 1975; Lycan 1987]. I remain neutral about this problem. What my proposal suggests is that when a mental state results from a 'merely' subpersonal-level mental process, one that does not correspond to any personal-level mental process, the state does not need a good evidential basis in order to have justificatory power. Cognitively unpenetrated perceptions and illusions, I think, involve merely subpersonal-level mental processes.

The distinction between personal- and subpersonal-level mental processes is different from that between deliberate and spontaneous mental processes, or between conscious and unconscious mental processes. Deliberateness or consciousness is not required for personal-level mental processes. Many belief-forming mental processes are quite spontaneous or even unconscious, but they are nonetheless attributed to the person. In the blizzard case, there is a folk-psychological explanation of Bob's imagining: his imaginative experience is triggered by his hope that a blizzard will arrive soon. The fact that there is such an explanation makes the posited mental process personal-level. Whether the process is deliberate or spontaneous, conscious or unconscious, does not make a difference to its personal-level status.

Why is the distinction between personal- and subpersonal-level mental processes epistemically important? I think that there are two main reasons, of which the first is that justification is ultimately an evaluation of a subject, rather than a subpersonal system of a subject, or a different subject. When we say that a mental state is justified, what we mean is that the subject is justified to possess it, and when we say that a mental state has justificatory power, what we mean is that the subject can justifiably employ it to support certain beliefs.

To see an example, compare a belief inference you yourself make with a belief inference someone else makes, but whose outcome is inserted into your mind (say, through

hypnosis or direct manipulation of your brain states). There is an asymmetry between these two cases. Whereas the appropriateness of the former inference matters to the justificatory status and power of your inferred belief, the appropriateness of the latter inference does not matter to the justificatory status and power of your externally inserted belief. One reasonable explanation is that the second inference is not a mental process you yourself go through, and that is why its quality should not make a difference to your justificatory situation with respect to the belief.

The same point also holds for mental states that result from merely subpersonal-level mental processes. These mental processes are ascribed to subpersonal systems, but not you. The quality of the processes should not make a difference to your justificatory situation with respect to the mental states that they generate.

The other reason why the distinction between personal- and subpersonal-level mental processes is epistemically important is that if we took subpersonal-level mental processes to affect our justificatory status then we would likely end up with disagreeable epistemological conclusions. As mentioned when discussing Siegel's version of inferentialism, the perceptual system might engage in mental processes that fall short of standards of logic and rationality. Imposing a requirement for a good evidential basis could lead to extensive scepticism about the justificatory power of perceptual experiences. Besides, as the Müller-Lyer illusion shows, the perceptual system can generate this illusion regardless of whether we, on a personal level, have prior evidence that the lines are of the same length. If the quality of a subpersonal-level mental process mattered to our justificatory situation, then an assessment of the former might also need to take account of our personal-level mental states. One may have to judge that in the Müller-Lyer illusion, the subpersonal-level mental process neglects an important

piece of evidence that we possess, a neglect that causes the resulting experience to lack justificatory power. But we normally take such illusory experiences to give us justification.[16]

Before applying this epistemological theory to cognitive penetration cases, I make two further clarifications. First, fabricationism does not propose that when a mental state results from a personal-level mental process, having a good evidential basis is sufficient for the state to be justified or to have justificatory power. It merely takes having a good evidential basis as a necessary condition.

Second, one might wonder whether Bob's imagining in the blizzard case could have justificatory power if he had a justified belief that it is snowing, which served as the basis for his imagining. Fabricationism does not preclude this possibility. Even if Bob's imagining had justificatory power, it would not be the same as directly seeing snow. The justificatory power might be inherited from his justified belief, and might not give him any new justification for believing that it is snowing. Moreover, information from different sources can have different evidential weight. Suppose that the weather app on my phone reports that it is snowing, but I don't see any snow. The evidential force of my perception obviously outweighs that of the weather app's report. Similarly, even if Bob's imagining had justificatory power, the justificatory power would be weaker than directly seeing snow.

Cognitive penetration normally refers to a situation in which two subjects have perceptual experiences with different contents and phenomenal character due to differences in the subjects' cognitive states, when the following factors are held fixed—the distal stimuli, external conditions, conditions of the subjects' sensory organs, and their focus and spatial attention. This definition does not spell out how cognitive penetration occurs. If, as argued

[16] Smithies [2019] makes the point that if subpersonal-level mental states mattered to our justificatory situation, then, when a subpersonal-level mental state contradicts to some of our beliefs, we would have to conclude that we are being logically inconsistent, which is also unacceptable.

in section 4, cognitive states contribute components to the subjects' perceptual experiences through imaginative processes, then the mental processes, just like Bob's imagining, can be attributed to the subjects, taken as whole persons, and hence are subject to the requirement for a good evidential basis.

The anger case, preformationism case, and banana case can be explained this way. Jill's fear contributes an anger component to her perceptual experience of Jack through an imaginative process, but her imagining lacks a good evidential basis, and so the relevant part of her experience counts as fabricated. It is epistemically downgraded with respect to believing that Jack has an angry face. The preformationists' unjustified background belief contributed an embryo component to their perceptual experience through an imaginative process. Given the unjustifiedness of the background belief, the embryo part of their experience counted as fabricated, and their experience was epistemically downgraded with respect to believing that the observed sperm cells contained embryos.

The account of the banana case would be similar to that of the preformationism case if the subjects' background belief were unjustified. However, suppose that they were justified in believing that bananas are typically yellow. Then the yellow component of their experience does not count as fabricated. How to explain the epistemic downgrade?

Because the subjects' imaginative contribution was based on their background belief, their overall experience might not provide any new justification for believing that the banana picture was yellow, even if the experience provided some new justification for believing that the banana picture was grey (due to the perceptual contribution). The overall experience provided less new justification for believing that the banana picture was yellowish-grey than would an experience that is cognitively unpenetrated while representing the same content. Also, suppose that the subjects' experience provided new justification for believing that the

banana picture was yellow. The point that information from different sources can have different evidential weight once again becomes relevant. The evidential force of the subjects' experience, which had contribution from their imagining, still seems to be outweighed by that of a cognitively unpenetrated yellowish-grey perceptual experience. This helps to explain our persistent intuition that, even in the banana case, cognitive penetration detracts from the justificatory power of the experience.

In section 4, I argued that apparently innocent or good cognitive penetration could be explained by another kind of imagining-perception interaction: namely, cognitively triggered imagining primes, and hence facilitates, the perception of the relevant external object. I find it less clear that, in the friend case and birdwatcher case, the perceptual experiences result from personal-level mental processes. For here imaginings seem to only play an assisting role. It is hard to ascribe any content or any phenomenal character of the subjects' perceptual experiences directly to their imaginings. Despite being cognitively penetrated, the perceptual experiences might not contain any fabricated content, and give the subjects no less justification.

6. Concluding Remarks

So far, I have offered and defended an alternative explanation of the epistemological implication of cognitive penetration—fabricationism. One important difference between this view and inferentialism is that the former does not require cognitively penetrated perceptual experiences to result from inferences in order to be able to evaluate their justificatory power. One example that I offered to demonstrate the inadequacy of the middle-ground approach under inferentialism was that your desire that there is a yellow banana causes you to view a

banana picture as yellowish-grey when it is entirely grey. While the causal influence of your desire on your perceptual experience might not be a genuine inference, which poses a problem for the middle-ground approach, fabricationism can easily explain the epistemic downgrade of your experience by appealing to the fact that it results from a personal-level mental process while lacking a good evidential basis.[17]

References

Berger, C. C. and H. H. Ehrsson 2013. Mental Imagery Changes Multisensory Perception, *Current Biology* 23/14: 1367-1372.

Berger, C. C. and H. H. Ehrsson 2014. The Fusion of Mental Imagery and Sensation in the Temporal Association Cortex, *The Journal of Neuroscience* 34/41: 13684-13692.

Delk, J. and S. Fillenbaum 1965. Differences in Perceived Colour as a Function of Characteristic Colour, *The American Journal of Psychology* 78/2: 290-293.

Dennett, D. C. 1969. *Content and Consciousness*, London: Routledge and Kegan Paul.

Drayson, Z. 2012. The Uses and Abuses of the Personal/Subpersonal Distinction, *Philosophical Perspectives* 26/1: 1-18.

Drayson, Z. 2014. The Personal/Subpersonal Distinction, *Philosophy Compass* 9/5: 338-346.

Fodor, J. 1975. *The Language of Thought*, Cambridge, MA: Harvard University Press.

Ghijsen, H. 2016. The Real Problem of Cognitive Penetration, *Philosophical Studies* 173/6: 1457-1475.

[17] I owe special thanks to Ali Boyle, Georgi Gardiner, Matt McGrath, Bence Nanay, Susanna Siegel, Nico Silins, and Ru Ye for valuable comments and discussion. I am grateful to audiences at Cornell University, the University of Minho, the University of California, Berkeley, the University of Antwerp, the Ruhr-University Bochum, the University of Leeds, East China Normal University, Lingnan University, and Nanjing University. Finally, thanks to several helpful anonymous referees.

Hansen, T., M. Olkkonen, S. Walter, and K. Gegenfurtner 2006. Memory Modulates Colour Appearance, *Nature Neuroscience* 9/11: 1367-1368.

Hart, W. D. 1988. *The Engines of the Soul*, Cambridge: Cambridge University Press.

Kind, A. 2016. Imagining Under Constraints, in *Knowledge Through Imagination*, ed. A. Kind and P. Kung, New York: Oxford University Press: 145-159.

Kind, A. 2018. How Imagination Gives Rise to Knowledge, in *Perceptual Imagination and Perceptual Memory*, ed. F. Macpherson and F. Dorsch, New York: Oxford University Press: 227-246.

Kosslyn, S. M. and W. L. Thompson 2003. When Is Early Visual Cortex Activated During Visual Mental Imagery? *Psychological Bulletin* 129/5: 723-746.

Kriegel, U. 2015. *The Varieties of Consciousness*, New York: Oxford University Press.

Langland-Hassan, P. 2015. Imaginative Attitudes, *Philosophy and Phenomenological Research* 90/3: 664-686.

Levin, D. T. and M. T. Banaji 2006. Distortions in the Perceived Lightness of Faces: The Role of Race Categories, *Journal of Experimental Psychology: General* 135/4: 501-512.

Long, R. 2018. How Wishful Seeing is Not Like Wishful Thinking, *Philosophical Studies* 175/6: 1401-1421.

Lycan, W. 1987. *Consciousness*, Cambridge, MA: MIT Press.

Lyons, J. 2011. Circularity, Reliability, and the Cognitive Penetrability of Perception, *Philosophical Issues* 21/1: 289-311.

Lyons, J. 2016. Inferentialism and Cognitive Penetration of Perception, *Episteme* 13/1: 1-28.

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

McGinn, C. 2004. *Mindsight: Image, Dream, Meaning*, Cambridge: Harvard University Press.

McGrath, M. 2013. Phenomenal Conservatism and Cognitive Penetration: the "Bad Basis" Counterexample, in *Seemings and Justification*, ed. C. Tucker, New York: Oxford University Press: 225-247.

Mole, C. 2015. Attention and Cognitive Penetration, in *The Cognitive Penetrability of Perception: New Philosophical Perspective*, ed. A. Raftopoulos and J. Zeimbekis, New York: Oxford University Press: 218-238.

Morrison, J. 2016. Perceptual Confidence, *Analytic Philosophy* 57/1: 15-48.

Munton, J. 2016. Visual Confidences and Direct Perceptual Justification, *Philosophical Topics* 44/2: 301-326.

Munton, J. 2019. Perceptual Skill and Social Structure, *Philosophy and Phenomenological Research* 99/1: 131-161.

Nanay, B. 2018. Multimodal Mental Imagery, *Cortex* 105: 125-134.

Nanay, B. manuscript. *Seeing Things You Don't See.*

Olkkonen, M., T. Hansen, and K. Gegenfurtner 2008. Colour Appearance of Familiar Objects: Effects of Object Shape, Texture, and Illumination Changes, *Journal of Vision* 8/5: 1–16.

Page, J. W., P. Duhamel, and M. Crognale 2011. ERP Evidence of Visualization at Early Stages of Visual Processing, *Brain and Cognition* 75/2: 141-146.

Pearson, J., C. W. G. Clifford, and F. Tong 2008. The Functional Impact of Mental Imagery on Conscious Perception, *Current Biology* 18/13: 982-986.

Pearson, J. and S. M. Kosslyn 2015. The heterogeneity of mental representation: Ending the imagery debate, *Proceedings of the National Academy of Sciences* 112/33: 10089-10092.

Pearson, J., R. L. Rademaker, and F. Tong 2011. Evaluating the Mind's Eye: the Metacognition of Visual Imagery, *Psychological Science* 22/12: 1535-1542.

Perky, C. W. 1910. An Experimental Study of Imagination, *American Journal of Psychology* 21/3: 422–452.

Reddy, L., N. Tsuchiya, and T. Serre 2010. Reading the Mind's Eye: Decoding Category Information During Mental Imagery, *Neuroimage* 50/2: 818-825.

Searle, J. 1983. *Intentionality: An Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press.

Segal, S. J. 1971. Processing of the Stimulus in Imagery and Perception, in *Imagery: Current Cognitive Approaches*, ed. S. J. Segal, New York: Academic Press: 69-100.

Segal, S. J. 1972. Assimilation of a Stimulus in the Construction of an Image: The Perky Effect Revisited, in *The Function and Nature of Imagery*, ed. P. W. Sheehan, New York: Academic Press: 203-230.

Siegel, S. 2012. Cognitive Penetrability and Perceptual Justification, *Nous* 46/2: 201-222.

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

Siegel, S. 2017. *The Rationality of Perception*, New York: Oxford University Press.

Siegel S. manuscript. Does Perceptual Experience Come in Degrees?

Silins, N. 2016. Cognitive Penetration and the Epistemology of Perception, *Philosophy Compass* 11/1: 24-42.

Smithies, D. 2019. *The Epistemic Role of Consciousness*, New York: Oxford University Press.

Spence, C. and O. Deroy 2013. Crossmodal Mental Imagery, in *Multisensory Imagery*, ed. S. Lacey and R. Lawson, New York: Springer: 157-184.

Stokes, D. 2018. Attention and the Cognitive Penetrability of Perception, *Australasian Journal of Philosophy* 96/2: 303-318.

Stokes, D. forthcoming. On Perceptual Expertise, *Mind and Language*.

Stokes, D. and S. Biggs 2014. The Dominance of the Visual, in *Perception and Its Modalities*, ed. D. Stokes, M. Matthen, and S. Biggs, New York: Oxford University Press: 350-378.

Stokes, M., R. Thompson, R. Cusack, and J. Duncan 2009. Top-Down Activation of Shape-Specific Population Codes in Visual Cortex During Mental Imagery, *Journal of Neuroscience* 29/5: 1565-1572.

Teng, L. 2016. Cognitive Penetration, Imagining, and the Downgrade Thesis, *Philosophical Topics* 44/2: 405-426.

Teng, L. 2018. Is Phenomenal Force Sufficient for Immediate Perceptual Justification? *Synthese* 195/2: 637-656.

Vance, J. 2014. Emotion and the New Epistemic Challenge from Cognitive Penetrability, *Philosophical Studies* 169/2: 257-283.

Watzl, S. 2017. *Structuring the Mind: The Nature of Attention and How It Shapes Consciousness*, New York: Oxford University Press.

Wu, W. 2017. Shaking Up the Mind's Ground Floor: The Cognitive Penetration of Visual Attention, *Journal of Philosophy* 114/1: 5-32.

Yablo, S. 1993. Is Conceivability a Guide to Possibility? *Philosophy and Phenomenological Research* 53/1: 1-42.