Cognitive Penetrability
Ch 3 of *Seemings and Justification*

Luca Moretti
University of Aberdeen
l.moretti@abdn.ac.uk

3. Cognitive penetrability

3.1 Introduction

In this chapter I introduce the thesis that perceptual appearances are cognitively penetrable and analyse cases made against phenomenal conservatism hinging on this thesis. In particular, I focus on objections coming from the externalist reliabilist camp and the internalist inferentialist camp. I conclude that cognitive penetrability doesn't yield lethal or substantive difficulties for phenomenal conservatism.

3.2 Characterizing cognitive penetrability

Preformationism was the biological theory that organisms are fully formed in miniature within the ovum or the spermatozoon. Some microscopists who endorsed preformationism claimed that they actually saw embryos in the sperm cells (cf. Siegel 2012). This historical datum has more than one possible explanation. Perhaps these biologists didn't actually have visual experiences of embryos; they mistakenly *described* the things that they saw as embryos, or their visual experiences as experiences of embryos. Another possibility is, however, that they had — literally — visual experiences of embryos. If the content of one's experience can be *cognitively penetrated* by one's beliefs, it isn't implausible that biologists who *believed in preformationism* actually had visual experiences of embryos in sperm cells.

The thesis of cognitive penetrability of perception holds — roughly — that the content of one's perceptual experiences or appearances can partly be determined *directly* by one's prior or concurrent cognitive states — for example, beliefs, conjectures, expectations, suspects, desires, hopes, moods, emotions, traits, decisions, accepted values, pursued goals, and attitudes or abilities acquired through experience or training. It isn't easy to clarify what 'directly' means in this characterization (cf. Tucker forthcoming). The intuition is that not just any kind of influence on one's perceptual appearances by one's cognitive states would qualify as cognitive penetration. A

cognitive state of mine might influence my perceptual appearance simply because it causes a change in the *spatial location* of my sense organs, so that I receive new perceptual stimuli. Imagine I intend to watch a painting. Accordingly, I turn my head towards the painting. My perceptual experience changes from representing, say, a wall to representing the painting. This kind of influence doesn't count as direct. Intuitively, in this case, there is an *indirect* or *external* link in the causal chain from my intention to my experiential state. So there is no cognitive penetration (cf. Stokes 2012).

People have tried to supply more precise characterizations of cognitive penetrability. (See for instance Pylyshyn 1999, Siegel 2012, Macpherson 2012, Stokes 2012, Wu 2013 and Raftopoulos 2019). This is a basic characterization that focuses on visual appearances. Take subjects S_1 and S_2, where S_1 and S_2 can be the same subject in different times or different counterfactual circumstances.

(CP) Visual appearances are cognitively penetrable just in case the following scenario is nomologically possible:

> When S_1's and S_2's organs of sight function equally well and are oriented towards the same distal stimulus, under the same external conditions, S_1 and S_2 have visual appearances with different contents because of S_1's and S_2's concurrent or antecedent cognitive states (cf. Macpherson 2012 and Siegel 2012).

Suppose the distal stimulus is an apple. In the scenario envisaged in (CP), S_1 and S_2 have organs of sight that work equally well. Since both subjects watch in the direction of the apple, they have the same sensory inputs. Furthermore, S_1 and S_2 are in the same external conditions. This entails, for instance, that the quantity and quality of light available to S_1 and S_2 is the same. Finally, S_1 and S_2

have visual appearances with different contents partly caused by their different cognitive states. If all these conditions are satisfied, visual appearances are cognitively penetrable.

(CP) comes together with a characterization of cognitively *penetrated* perceptual appearance. Suppose that in the above scenario, S_1 has a cognitive state that S_2 lacks. For instance, only S_1 but not S_2 has the *expectation* that the apple is green. If S_1's visual seeming is different from S_2's — e.g. only S_1's experience represents the apple as green — because of S_1's expectation, this state *penetrates* (the content of) S_1's appearance.

To familiarize with the notions introduced, let's consider two imaginary cases of cognitive penetration often discussed in the literature.

Angry Jack

Jill believes for no good reason that Jack is angry. This belief (or the arising expectation) affects the way Jack looks to her, producing in her a visual seeming that Jack is angry. Jill thus believes, even more firmly than before, that Jack is angry (cf. Siegel 2012).

Prospectors

Gus and Virgil are gold prospectors. Gus is an expert. He has learned to do so through training and experience. He began with a list of identification criteria and consciously applied them. He then reached the point where he could "see" that a nugget is gold. Virgil is a novice. He has a general sense of what gold looks like, but he isn't very good at its visual identification. While looking for gold, they find a pebble. Gus has the visual appearance that the pebble is gold as a result of his expertise. So he believes that the little stone is gold. Virgil also has a visual appearance that the pebble is gold, but only as a result of his desire of being rich. Like Gus, Virgil comes to believe that the pebble is gold (cf. Markie 2013).

Consider again (CP). In *Angry Jack* the distal stimulus is Jack's face, S_1 corresponds to Jill, and S_2 corresponds to Jill in a counterfactual situation in which she watches Jack's face in the same external conditions but she doesn't have the belief that (*P*) Jack is angry.[1] Jill's belief that *P* is involved in causing Jill's appearance that *P*. In the counterfactual situation, however, Jill doesn't have the belief that *P* and, consequently, she doesn't have the appearance that *P*. If all this is nomologically possible, in accordance with (CP), experience is cognitively penetrable. This happens because Jill's seeming that *P* can be penetrated by her belief that *P*.

Take now *Prospectors.* In this case the distal stimulus is the pebble. Let's first identify S_1 with Virgil, and S_2 with Virgil in a counterfactual situation in which he inspects the pebble under the same external conditions but he doesn't crave for becoming rich. Virgil's desire to become rich is involved in causing his seeming that (*Q*) the pebble id gold. In the counterfactual situation, however, Virgil has no desire to be rich[2] and, accordingly, he doesn't experience that *Q*. If all this is nomologically possible, in accordance with (CP), perceptual appearances are cognitively penetrable. This happens because Virgil's desire to become rich can penetrate his experience that *Q*.

Let's now identify S_1 with Gus, and S_2 with Gus in a counterfactual situation in which he still inspects the pebble under the same external conditions but he has no perceptual expertise. Gus' perceptual expertise is involved in causing his seeming that *Q*. In the counterfactual situation, however, Gus has no perceptual expertise and thus no appearance that *Q*. If all this is nomologically possible, appearances are cognitively penetrable, for Gus' perceptual expertise can penetrate his appearance that *Q*.

As we have seen in §2.3, many philosophers think that any perceptual appearance is a compound of two elements: a sensory state and a seeming (where the seeming is a sort of interpretation of the sensory state). If this conception is correct, there are two prima facie possible

1 Or in which she has a belief that *P* that has no causal bearing on her seemings.
2 Or he has a desire to be rich that has no causal bearing on his seemings.

ways in which a subject *S* might come to have a perceptual seeming that *P* cognitively penetrated by a state *M* of *S*. *M* might directly cause or change *a sensory state* of *S*, and this might in turn produce an appropriate *seeming* that *P* in *S*. (For instance, under the influence of her belief that Jack is angry, Jill could have a sensory impression of two slanted eyes and a frowning mouth while watching Jack's face. This sensory state could then cause Jill's seeming that Jack is angry.) Alternatively, *M* might directly cause the seeming that *P* in *S* without affecting *S*'s sensory states. (For instance, Jill might have the visual sensation of a non-angry face, but her belief that Jack is angry could directly produce her seeming that Jack is so.) In this case, there might be a mismatch detectable by *S* between her sensory states and *S*'s interpretation of them (cf. Huemer 2013b).

If the *monistic* conception of perceptual seemings proved correct (see above §2.3), the contents of perceptual seemings might include representations of low-level properties of things (such as shapes and colours) and representations of high-level properties of things (such as being a dog or being a computer). In this case too there are two prima facie possible ways in which *S* might come to have a perceptual seeming that *P* cognitively penetrated by a mental state *M*. *M* might directly produce a representation of low-level properties in *S* that in turn elicits a representation of high-level properties, or *M* might directly cause a representation of high-level properties in *S*.

I have characterized cognitive penetrability in terms of (CP), but the adequacy of this characterization can be questioned. Philosophers could contend that (CP) doesn't discriminate between the genuine cognitive penetrability of one's perception and the ability of one's *covert* attention to select some features of the distal stimulus as opposed to some others. *S*'s attention to something *X* (in the external environment) is *overt* if it is implemented by *S*'s *bodily* behaviour — e.g. by *S*'s pointing her eyes in the direction of *X*. *S*'s attention to *X* is *covert* if its implementation only involves a selection process *internal* to *S* that makes no bodily difference (cf. Mole 2015). Suppose that S_1's and S_2's organs of sight function equally well and are oriented towards the same distal stimulus *X*. Imagine that under the same external conditions, S_1 and S_2 have visual seemings

with different contents because of S_1's and S_2's previous cognitive states make them *covertly attend to different features of X*. If *X* is an apple, S_1 might focus on its shape, and S_2 on its colour. On (CP), we should conclude that perceptual appearances are cognitively penetrable. Yet many would deny this. Many would insist that a mental state can produce cognitive penetration only if it affects *S*'s *perceptual processing itself*. If a state of *S* causes a change in the content of *S*'s appearance only because it determines the focus of *S*'s covert attention *before she has the experience*, there seem to be at best an *indirect* link from *S*'s cognitive state to *S*'s appearance. This doesn't look like cognitive penetration (cf. Siegel 2012, 2013a, Silins 2016 and Tucker forthcoming).[3]

To settle this difficulty, we might supplant (CP) with a principle of this sort:

(CP*) Visual appearances are cognitively penetrable just in case the following scenario is nomologically possible:

> When S_1's and S_2's organs of sight function equally well and are oriented towards the same distal stimulus *X*, and S_1 and S_2 *covertly attend to the same the same features of X,* under the same external conditions, S_1 and S_2 have visual appearances with different contents because of S_1's and S_2's concurrent or antecedent cognitive states (cf. Siegel 2012 and Silins 2016).

A problematic aspect of the notion of covert attention presupposed in (CP*) and in our discussion so far is that it assumes that this type of attention is pre-experiential. Covert attention is thought of as a process external to the relevant experience, for the allocation of attention to a spatial location is taken to occur *before* the experience occurs. Although such a conception seems to be

3 The phenomenon of selection by covert attention (as opposed to cognitive penetration) *might* have by itself problematic implications for phenomenal conservatism, which I won't investigate in this work. On this issue see Tucker (forthcoming).

taken for granted by some authors — e.g. Pylyshyn (1999), Siegel (2012, 2013a, 2017) and Firestone and Scholl (2016), other scholar explicitly reject it — e.g. Mole (2015) and Wu (2017). The latter describe findings of psychology and cognitive sciences that strongly suggest that the selection processes of covert attention are *constitutive elements* of the relevant appearances. If this is correct, the selective ability of covert attention may count as a type of cognitive penetrability, and (CP*) should be dropped as inadequate.[4]

Because of these complications, in the remainder of this chapter I will simply assume that some principle *in the neighbourhood of (CP) and (CP*)* does provide an appropriate characterization of cognitive penetrability. Whatever specific characterization is adopted, it won't affect the adequacy of my analyses and arguments in the next sections.

3.3 The epistemic problem of cognitive penetrability

Although the cognitive penetrability of perception remains to date a controversial empirical hypothesis, it has often been adduced to challenge the views that hold that perceptual experiences have inherent justifying power. Phenomenal conservatism has been the target *par excellence* (see for instance Lyons 2011, Siegel 2012, Brogaard 2013, Markie 2013, McGrath 2013a and Teng 2016).

Those who adduce examples like *Prospectors* and *Angry Jack* to challenge phenomenal conservatism argue that it is intuitive that in these situations Jill and Virgil cannot have justification for holding their perceptual beliefs. So Virgil's seeming that the pebble is gold would be unable to give him justification for believing that the pebble is gold, and Jill's seeming that Jack is angry would be unable to give her justification for believing that Jack is angry.[5] As we will see, there are

4 For opinionated overviews of this multifaceted debate see for instance Gatzia and Brogaard (2017) and Raftopoulos (2019).

5 In a variant of *Angry Jack*, Jill starts with a *justified* belief that (*P*) Jack is angry, which causes in Jill a seeming that *P*. Siegel (2012) adduces this variant to contend that phenomenal conservatism is false or implausible. The argument is that Jill's justification coming from her appearance that *P* would need to *boost* Jill's original justification for believing *P*, which is intuitively impossible in this case. So Jill cannot have justification from her appearance. The central

various ways to unpack these intuitions. The epistemologists who challenge phenomenal conservatism by adducing examples of these types insist that the perceptual seemings of the subjects in these situations cannot justify their contents *whether or not the subjects are aware that the seemings are cognitively penetrated*. So these examples are not thought of as cases of defeat. The contention is that the subjects in question cannot acquire even *prima facie* justification for their beliefs from the cognitively penetrated seemings (cf. Lyons 2011, Markie 2013, McGrath 2013a and Siegel 2013b). This thesis is incompatible with (PC). As we have seen in§2.2, (PC) entails that if it seems to *S* that *P*, *S* thereby has prima facie justification for believing *P*.[6]

Not just all cases of cognitive penetration (real or supposed) are deemed to be cases of *bad* cognitive penetration — i.e. situations in which the subject has no prima facie justification for the relevant belief. Some think that there are cases of *good* cognitive penetration — i.e. situations in which the subject does have prima facie justification (or *more* prima facie justification) for a perceptual belief just because she has cognitively penetrated seemings. *Perceptual learning* is often deemed to produce good cognitive penetration. Perceptual learning is a process based on training and experience that ends up causing changes in the subject's perceptual-discriminating abilities. An example could concern the cognitively penetrated visual appearances of a trained radiologists who inspects X-rays to detect a bone fracture (cf. Siegel 2012). Another example could be Gus' cognitively penetrated seeming that the pebble is gold in *Prospectors*. Some epistemologists suggest that there may be cases of good cognitive penetration independent of perceptual learning. Lyons (2011) imagines a scenario in which a subject — let's call him John — gets his snake-detection skills sharpened up because his perceptual appearances are penetrated by his fear that there are

problem of this argument is that it presupposes that a justification based on an appearance can boost a justification already available for a belief even if the subject ignores whether the appearance is independent of the belief. (Jill, for instance, ignores this.) But it is dubious or unclear that the phenomenal conservative is committed to this principle (cf. Tucker 2013 and Huemer 2013a).

6 Although the current dispute mainly focuses on *visual* appearances, cognitive penetrability might affect other types of perceptual appearance. It is also in principle conceivable that phenomena analogous to cognitive penetrability could affect, for example, mnemonic, rational or moral seemings.

snakes around. Suppose for example that, due to his fear, all snake-shapes *stand out* in John's visual experience, so that he can see the reptiles despite their camouflages.[7] The authors who appeal to cognitive penetrability to challenge phenomenal conservatism primarily adduce alleged counterexamples based on cases of bad cognitive penetration. Some also insist that phenomenal conservatism is problematic because it cannot spell out the opposite epistemic consequences of bad and good cognitive penetration (cf. Lyons 2011).

The objections to phenomenal conservatism that adduce cognitive penetrability come from both the externalist and the internalist camp. That internalists level objections of this type is remarkable. There is in fact a striking similarity between a perceptual appearance that *S* would have if *S* were in a sceptical scenario (such as the Cartesian demon scenario or the Matrix scenario), and an appearance of *S* cognitively penetrated. Both these appearances would have *anomalous aetiologies*. In the first case, the aetiology would be anomalous because the distal cause of *S*'s appearance would be *unnatural* (e.g. if *S*'s visual experience of a cat were caused by a demon, the distal cause of *S*'s appearance would be the demon rather than a cat). In the second case, the aetiology would be anomalous because a mental state of *S* would interfere with *S*'s *normal* causal chains that produce perceptual appearances (e.g. in *Angry Jack*, Jill's belief that Jack is angry interferes with Jill's normal visual processes). Internalists generally agree that if *S* were in a sceptical scenario, the anomalous aetiologies of *S*'s seemings would *not* affect the justifying power of these appearances. A frequently adduced reason is that the aetiologies wouldn't be reflectively accessible to *S*. One might thus expect that internalists also agree that if *S*'s appearances were cognitively penetrated, their anomalous aetiologies wouldn't affect their justifying power. For these aetiologies (at least segments of them) wouldn't be reflectively accessible to *S*. However, certain self-avowed internalists — called by Lyons (2016) *inferentialists* — insist that if *S*'s seemings were

7 Note that this might be argued to be a mere *attentional* effect unrelated to cognitive penetration (cf. Siegel 2017: 123-125).

badly cognitively penetrated, the flawed aetiologies of these appearances would undermine their justifying power. Since these internalists conceive of the aetiologies of perceptual appearances as sequences of *mental states,* they are committed to embracing a *mentalist* variant of internalism (cf. Lyons 2016).

In the next two sections, I criticize externalist and internalist cases against phenomenal conservatism that appeal to cognitive penetrability. I will dedicate much more attention and space to internalist arguments because — as clarified in §1 — in this work I presuppose that internalism is correct.[8] Before starting off, let me call attention to a crucial premise of this debate. As said, the cognitive penetrability of perceptual appearances is a *controversial empirical hypothesis*. Even so, phenomenal conservatives couldn't hope to dismiss objections that appeal to cognitive penetrability by simply emphasizing that it is an unattested phenomenon. The problem is that phenomenal conservatives conceive of (PC) as an *a priori* true principle (cf. Huemer 2007). Hence, even if our actual hardwiring ruled out cognitive penetrability, the mere *conceptual possibility* of a rational subject whose seemings lack justifying power because they are cognitively penetrated would be incompatible with (PC) (cf. Markie 2013, Georgakakis and Moretti 2019 and Tucker forthcoming).

3.4 Reliabilism, cognitive penetrability, and phenomenal conservatism

Externalists typically explain the asserted intuition that cognitive penetration alters the justifying power of appearances by invoking reliabilism. They contend that the malformed aetiologies of cognitively penetrated appearances make the belief production processes based on those appearances unreliable. Since reliabilism in essence equates the justification of a belief with the reliability of the process that has produced it, the conclusion is that beliefs based on cognitively penetrated appearances are unjustified. More accurately, externalist reliabilists contend that *bad*

8 In epistemology cognitive penetrability has also been approached from perspectives that don't fit well the dichotomy internalist/externalism. (I include Brogaard 2013 in this group. Brogaard describes herself as an epistemic *internalist* but the view she defends — sensible dogmatism — has a marked *reliabilist* component.) For lack of space I cannot analyse these interesting accounts here. For an overview see Georgakakis and Moretti (2019).

cognitive penetration affects reliability — and thus justification — *negatively*, by destroying or reducing it, whereas *good* cognitive penetration affects reliability — and thus justification — *positively*, by generating or enhancing it (see mainly Lyons 2011, 2016 and Ghijsen 2016).[9]

For example, reliabilists would maintain that the reason why it is counterintuitive that, in *Angry Jack,* Jill's perceptual belief is prima facie justified is the following: the belief production process deployed by Jill is unreliable because it encompasses the initial ungrounded belief of Jill that Jack is angry. Reliabilists would also maintain that the reason why it is intuitive that in *Prospectors* Gus' perceptual belief is prima facie justified while Virgil's identical belief is not is the following: the belief production process used by Gus is reliable, while the one used by Virgil is unreliable. Gus' belief production process in fact exploits detection skills that Gus is endowed with, whereas Virgil's belief production process is ultimately based on a mere desire of Virgil. Cases of good cognitive penetration independent of perceptual learning — such the snake-detection one — could be handled likewise. Reliabilists could claim that in the situation imagined by Lyons, John's perceptual beliefs that there are snakes in the trail are prima facie justified because John's fear has sharpened up his snake-detection skills, and so enhanced the reliability of the processes that yield the correlated beliefs.

The reliabilist contends that phenomenal conservatism cannot explain the intuition that cognitive penetration affects negatively or positively the ability of appearances to supply prima facie justification. For these effects of cognitive penetration essentially rest — according to the reliabilist — on reflectively *inaccessible* features of seeming production processes (i.e. their reliability), whereas (PC) makes prima facie justification depend on reflectively *accessible* factors only (see mainly Lyons 2011). If this is true, the adequacy of phenomenal conservatism is in peril.

9 There are two basic forms of reliabilism about justification: *process* reliabilism (outlined in §2.2) and *indicator* reliabilism. The latter states that *S*'s belief that *P* is prima facie justified just in case *S* bases it on a mental state *M* that reliably indicates *P*; where *M* reliably indicates *P* just in case in most of the closest possible worlds in which a subject has *M*, *P* is true. I focus on process reliabilism but my claims can be re-phrased to apply to indicator reliabilism.

But phenomenal conservatives have good counterarguments at hand. Firstly, they can put reliabilists under pressure by directly attacking the adequacy of the reliabilist conception of epistemic justification, which is notoriously afflicted by longstanding problems such as the *new evil demon* and the *generality* one (see Goldman and Beddor 2016 for useful discussion). More interestingly for this chapter's theme, phenomenal conservatives can adduce thought experiments to show that our epistemic intuitions about good and bad cognitive penetration don't get really explained or clarified by considerations about the reliability of one's belief production processes.

Take again Virgil from *Prospectors* but consider this variant of it. Virgil is unable to recognize the distinctive features of gold. However, it is true that whenever he entertains an appearance that a pebble is gold cognitively penetrated by his desire to become rich, very often the pebble *is* gold. Suppose that a benevolent demon incessantly intervenes to make this possible (cf. Fumerton 2006: 80). In this new scenario, Virgil's perceptual belief production process is *reliable* when Virgil's belief is based on an appearance that an object is gold cognitively penetrated by his desire to be rich. Suppose Virgil has one of these appearances and bases a belief that the object is gold on it. Many would say that Virgil's belief still looks *epistemically defective* despite the belief production process used by Virgil is reliable. This suggests that the reliabilist account of our intuitions about the epistemic consequences of bad cognitive penetration misses the target.

One can criticize the reliabilist account of our intuitions about the different epistemic consequences of good and bad cognitive penetration by exploiting Markie (2013)'s variant of *Prospectors*.[10] Suppose again Gus is an expert prospector and Virgil a novice. Unbeknownst to them, however, they have always been living in the Matrix. Imagine both Gus and Virgil have a perceptual seeming that a given nugget is gold. Gus' appearance is caused by the external stimuli due to the Matrix[11] and partly by his expertise and background knowledge. This looks like a case of

10 See also McGrath (2013b).
11 These stimuli are identical to those that his brain would receive in the real world.

good cognitive penetration. Virgil's seeming is caused by the same external stimuli from the Matrix and partly by his desire to become rich. This looks like a case of *bad* cognitive penetration. Imagine that both Gus and Virgil come to believe that the nugget is gold on the grounds of their seemings. The epistemic standing of Virgil's belief is *intuitively worse* than the epistemic standing of Gus' belief. But — note — the perceptual belief production processes used by *both* subjects are now completely unreliable. So reliability cannot be adduced to explain this difference. This example suggests that the reliabilist account of our intuitions about the divergent epistemic consequences of bad and good cognitive penetration is flawed.[12]

In conclusion, the reliabilist explanation of our epistemic intuitions about cognitive penetrability — at least in its basic form — appears flawed. Some reliabilists have put forward sophisticated variants of their basic conception of epistemic justification in which the link between one's justification and the reliability of one's mental processes is more indirect and complex (see for instance Goldman 1979, Comesaña 2002, 2010 and Bergmann 2006). Reliabilists might turn to these views to try to illuminate the epistemic intuitions about cognitive penetrability elicited by thought experiments like those above. A concern is, however, that these more sophisticated conceptions of justification would produce more complicated and less immediate and straightforward explanations of those intuitions, which ultimately wouldn't be preferable to internalist explanations. For more specific objections see Tucker (2014a).[13]

3.5 Inferentialism, cognitive penetrability, and phenomenal conservatism

Let's turn to the family of internalist mentalist views that Lyons (2016) calls *inferentialism* — namely, Siegel (2012, 2013b, 2017)'s *process inferentialism,* and McGrath (2013a, 2013b)'s and Markie (2013)'s versions of *evidence inferentialism*. Inferentialism aims to account for our

12 The first and the second thought experiment together suggest that unreliability is neither a sufficient nor a necessary condition for bad cognitive penetration.
13 See also Tucker (2014b), Vahid (2014) and Siegel (2017: §6).

epistemic intuitions about cognitive penetrability by appealing to the *rationality* of the subject's cognitive processes rather than their reliability. Its pivotal assumption is that when *S* has a cognitively penetrated appearance, the appearance always or typically results from *S*'s *doing.*[14] A consequence is that the epistemic standing of this appearance and its justifying power can be appraised by assessing whether *S* has produced it in a rational way. This can be done in roughly the same way in which we appraise the epistemic standing and the justifying power of a *belief* of *S* by assessing how *S* has *inferred* it from other beliefs (cf. Lyons 2016). Thus, whether a cognitively penetrated appearance can justify believing its content hinges on whether the appearance's aetiology is epistemically rational. The inferentialist opposes phenomenal conservatism by insisting that since (PC) doesn't make the justifying power of appearances depend on their aetiologies, phenomenal conservatism cannot explain the epistemic bearing of cognitive penetrability.

Inferentialism isn't vulnerable to the objections I have raised against the reliabilist account of cognitive penetrability. For the rationality of the aetiologies of perceptual appearances (supposing they can be rational) is logically independent of their reliability. For instance, the inferentialist could claim that Virgil's belief that his pebble is gold based on a badly cognitively penetrated appearance would be unjustified because its aetiology is irrational even if the appearance were yielded by an actually reliable process thanks to a benevolent demon's intervention.

Process inferentialism

Siegel (2017) offers the most developed inferentialist account to date. She accepts a very liberal notion of inference, according to which inferring is just a distinctive type of responding to an informational state that produces a conclusion epistemically dependent on that state. The informational state may be unconscious and can consist of a belief (or a similar doxastic state), an

14 Perhaps Markie would claim that this assumption is true only when cognitive penetration is good, and that bad cognitive penetration typically don't result from *S*'s doing.

appearance, a fear, a desire, or a combination of these elements. The conclusion can be a doxastic state or an appearance. The informational state and the response may occur simultaneously. Finally, the process of inferring doesn't have to feel like anything, it need not leave any mark in consciousness (cf. §§2-8).

Siegel doesn't analyse any further this general characterization. Yet she provides examples to distinguish this type of transition from those that *fail* to respond informational states (e.g. bypassing evidence), those that respond to *non-informational* states (e.g. associating thoughts), and those that respond *non-inferentially* to informational states (e.g. directing one's attention) (cf. §5).[15]

Let's examine how so a conceived inference performed by a subject *S* would bear on the epistemic standing of both its conclusion and *S* herself. Siegel dubs *epistemic charge* the property of a mental state of *S* in virtue of which the state and *S* are epistemically appraisable (cf. §§2-3). Epistemic charge is a property more general than epistemic justification. In the case of belief, justification and epistemic charge go hand-in-hand; no belief has one of these unless it has both. But mental states that cannot be justified — like appearances — can also have epistemic charge. Like justification, epistemic charge can be transmitted from a mental state to another. Depending on the *sign* of the charge, a mental state is enabled to justify other mental states. If a seeming that *P* of *S* has a *positive* charge, *S* and the seeming enjoy a positive epistemic status, so the seeming can prima facie justify *S*'s believing *P*. Conversely, if *S*'s seeming that *P* has a *negative* charge, *S*'s and the seeming's epistemic statuses are downgraded, so the seeming cannot even prima facie justify *S*'s believing *P*. The epistemic charge of a mental state can be *modulated* by the rationality of mental process that has produced the state. In particular, the epistemic charge of an appearance can be modulated by the rationality of the inference drawn by *S* that yields it.[16] A *good* inference concluding with a seeming bestows on it a *positive* charge. A *poor* inference concluding with a

15 Siegel doesn't consider Helmholtz-style "unconscious inferences" to be proper inferences, for these transitions are typically conceived of as mere *causal* processes that don't redound on the subject's rational standing (cf. Siegel 2019).
16 Inferences modulate but don't *generate* the epistemic charge of appearances. Siegel takes the epistemic charge of appearances to stem from what I have called, in §2.3, their *phenomenal force* (cf. 2017: 43–47).

seeming bestows on it a *negative* charge (cf. §§2, 4 and 6-7).[17] For Siegel, good cognitive penetration results from a good inference by *S*, and bad cognitive penetration results from a bad inference by *S*. When *S*'s experiences are badly cognitively penetrated, *S*'s overall outlook on the world sustains itself circularly; in the sense that it creates illusory appearances that represent the world in the way the outlook characterizes it. This is why Siegel says that badly cognitively penetrated appearances are *hijacked* by the subject's outlook (cf. §1).[18]

Siegel (2017: §§6 and 8) offers some examples of poor inferences ending with appearances that would instantiate epistemic flaws such as *circularity*, *inherited inappropriateness* or *jumping to conclusions*. For ease of illustration, let's dwell on an example of the last type. Suppose you infer a mental state *B* from a mental state *A,* but that the content of *A* doesn't support the content of *B*. This poor inference can downgrade the epistemic status of *B* and disable its justifying power. All this is uncontroversial if *A* and *B* are beliefs. Siegel contends that jumping to conclusions can also be instantiated by an inference *from a perceptual appearance to another perceptual appearance*. An example would be a variant of *Angry Jack* in which Jill entertains an appearance *B* that Jack is angry by jumping to conclusions from her appearance *A* that Jack has a neutral (non-angry) face, due to Jill's unfounded fear that Jack is angry. Siegel concedes that since *A* and *B* would occur in Jill's mind simultaneously, what Jill's fear brings about is not a *temporal transition* from *A* to *B* but a *dependence relation* of *B* on *A* (such that, presumably, *B* exists in virtue of *A*).[19] Thanks to this dependence relation, *B* would qualify as a response by Jill to the informational state *A*. So Jill can be described as inferring *B* from *A*. Since the content of *A* doesn't actually support the content of *B*, this is a bad inference that downgrades *B*'s epistemic status and makes *B* unable to justify Jill's belief that Jack is angry (cf. 117-119).

---

17 For Siegel there are various ways in which experiences can be subject to rationality — being the conclusion of an inference is *one* of them (cf. 2017: 21).

18 Siegel includes among the hijacked appearances also those *non-cognitively penetrated* that arise from *S*'s biased selection processes.

19 From now on, I will use 'transition' to refer to atemporal dependence relations of this type too.

In relation to this example, Siegel makes two additional claims. She suggests that a subject (e.g. Jill in another variant of *Angry Jack*) could have a perceptual appearance as a result of jumping to conclusions from an *unconscious perception* (cf. 123, see also 101-103 and 114). Siegel also contends that fears and desires could legitimately work as inputs to inferences (like beliefs and perceptual appearances), rather than counting as "disturbing factors" (cf. 123 and §8).

Although original and interesting, Siegel's account of cognitive penetrability is troubled by serious difficulties. One turns on Siegel's claim that to infer a state *B* form a state *A*, *S doesn't* need to take *A* to support *B* — where *S* takes *A* to support *B* just in case *S* appreciates (in some non-factive sense of this term) that the content of *A* is evidence for the content of *B* (cf. 2017: 95). Siegel's claim is worrisome because many if not most philosophers have long accepted the *taking condition*:

(TC) *S infers B* from *A* only if *S* comes to entertain *B because of* her taking *A* to support *B*.

See for instance Locke (1689/2008: book IV, Ch. 17.2), Frege (1979), Pierce (1905), Russell (1920), Stroud (1979), Broome (2014), and Boghossian (forthcoming).

There are various reasons why (TC) looks plausible. To begin with, it helps us produce a straightforward explanation of why inferring is more than associating thoughts: when we merely associate two thoughts, we don't take one to support the other. We do this when we infer a thought from another. Siegel has an alternative explanation. She proposes that when we associate two thoughts, we just respond to *concepts* involved in the first thought. In other words, the second thought is triggered only by concepts, and not by *predicatively-structured* (or truth-evaluable) segments of the first thought. (Imagine for example that while observing at dusk that the sky is growing dark, I recall that I need to buy lightbulbs. For Siegel, this would be a mere association of thoughts elicited by, say, my linking the concept of darkness with the concept of light.) On the other

hand, when we genuinely infer a thought from another — according to Siegel — we respond to predicatively-structured segments of the first thought (cf. 2017: 87). I don't think Siegel's proposal clarifies the difference between associating and inferring. Imagine for example that my thought that

(*P*) Mickey Mouse has a dog

draws my attention to the fact he has no cat, and so it triggers my thought that

(*Q*) Mickey Mouse has no cat.

Here it is my thought that *P as a whole* that triggers my thought that *Q*. So I'm responding to the *predicatively-structured* information that *P* by entertaining *Q.* This doesn't look like an inference, though.

(CT) also supplies a clear answer to the question why inferring is, intuitively, a form of *epistemic agency*. If (CT) is satisfied, it is *the agent* who carries out the inference once she appreciates that the premises support the conclusion. This appeal to the agent also helps us illuminate why inferences are more or less rational. Inferences are so because the agent't doing can be more or less epistemically appropriate.[20] One could try to reject the intuition that inferring is a form of epistemic agency or the intuition that the rationality of inference arises from the agent's doing. Yet Siegel grounds her explanation of cognitive penetrability in both intuitions. Since Siegel rejects (CT), it is unclear how her explanation could hope to account for these intuitions.

Appealing to (CT) also offers a neat explanation of what distinguishes an *arational* mental transition from a *poor* inference: while an arational transition doesn't satisfy (CT), a poor inference does. In a poor inference, the subject carries out the transition in virtue of her appreciating, *mistakenly,* that the premises support the conclusion. One might try to deny that there is a substantive difference between arational transitions and poor inferences. But Siegel couldn't. For she explicates bad cognitive penetration in terms of poor inferences, which she tells apart from

20 Interestingly, Siegel (2019) countenances that a subject *S* might somehow fulfil (CT) at *subpersonal* level. However, she acknowledges that this not the standard view, for the fulfilment of (CT), in this case, could no longer be adduced to explain the rationality of inference as resulting from *S*'s doing.

arational transitions. Since she rejects (CT), it is unclear why, in her account, arational mental transitions are substantially different from poor inferences.[21]

Some philosophers think that (CT) is implausible because it involves overintellectualisation. These philosophers observe that small children and perhaps higher animals can draw inference though they lack conceptual resources to grasp notions — like that of reason, rational support, premise and conclusion — which the subject must possess, according to these authors, to comply with (CT) (cf. McHugh and Way 2016). Another criticism states that (CT) cannot be satisfied because attempting to do so would elicit a vicious infinite regress (cf. Boghossian 2014 and McHugh and Way 2016). My view is that these objections miss the target, for they interpret the claim that *S* should appreciate that the content of *A* is evidence for the content of *B* as stating that, substantially, *S* should *rationally believe or accept* that the first content supports the second. But there is no need for such a demanding reading. These difficulties fade away when this claim is interpreted as the more modest proposition that it should *seem* to *S* that the content of *B* is likely or plausible in light of the truth of *A*.[22]

Another possible objection to (CT) comes from the observation that inferences we routinely draw don't seem to satisfy it. As Boghossian (forthcoming) suggests, however, phenomenology can be deceptive. For instance, non-occurrent beliefs are an important part of our mental and rational life, though they come without any phenomenology. We become aware of these beliefs only if we are stimulated with questions or external circumstances. I think that the apparent unfulfilment of (CT) can in most cases be explained along similar lines. When we genuinely draw an inference, irrespective of the phenomenology, if asked why we have come to believe the conclusion, we would normally realize that we have done so because we have taken the premises to support it.[23]

---

21 See Boghossian forthcoming for further reasons in support of (CT).
22 McHugh and Way (2016: 319-320) insist that this won't solve any problem. But their arguments are just sketched. My impression is that the view presented in Huemer (2016) isn't affected by the difficulties they allude to.
23 I wrote 'normally' because some psychological findings suggest that in certain cases, when people think they are describing the bases of their beliefs and decisions, they are instead merely theorizing or confabulating (see for instance Johansson et. al. 2006).

Siegel attempts to reject (CT) by adducing examples of genuine inferences that would not satisfy this constraint. Let's consider one of them. (Siegel's cases aren't essentially dissimilar, so by criticizing one, I'll criticize all of them.)

Kindness

> The person ahead of you in line at the Post Office is finding out from the clerk about the costs of sending a package. Their exchange of information is interspersed with comments about recent changes in the postal service and the most popular stamps. As you listen you are struck with the thought that the clerk is kind. You could not identify what it is about the clerk that leads you to this thought. Nor could you identify any generalizations that link these cues to kindness. Though you don't know it, you are responding to a combination of what she says to the customer, her forthright and friendly manner, her facial expressions, her tone of voice, and the way she handles the packages. (2017: 95)

In *Kindness* you are meant to reach a conclusion *B* — your belief that the clerk is kind — from a specific set *A* of cues that you possess. Siegel insists that this is a genuine inference (cf. 96). She stresses that in the imagined circumstances, nevertheless, you cannot identify the *A*-cues you are responding to. So you don't believe that they indicate kindness. Nor do you have a seeming or an intuition that these *A*-cues indicate kindness. Finally, you have no disposition to judge that these *A*-cues are cues for kindness. You might be disposed to say 'people who act *like that* are kind', but you cannot understand what '*like that*' picks out. Siegel concludes that in *Kindness,* you infer that the clerk is kind but you don't fulfil (CT). For 'there is no *X* such that [you] know [you are] responding to *X* in forming [your] belief' (97).[24]

24 Siegel (2019) returns to *Kindness* but her conclusion remains unchanged.

I think *Kindness* offers no evidence to believe that inferring doesn't require satisfying (CT). The basic problem is that the transition described in *Kindness* doesn't look like a process that most of us would name 'inference'. If one doesn't already presuppose a very liberal conception of inference that rejects (CT) — like Siegel's — one won't consider this transition to be an inference. Many would rather claim that *Kindness* depicts a typical case of *non-inferential* transition: one from a *seeming* to a belief. Indeed, it is natural to think that in *Kindness*, you would come to believe that the clerk is kind because it would *seem* to you that the clerk is so.[25] (You could have such a seeming even if you were unable to individuate specific features indicating that the clerk is kind.) To forestall criticism of this type, Siegel (2019) stipulates that, in *Kindness*, the clerk's kindness isn't *represented* in your experience. So 'you are struck with the thought that the clerk is kind' though it doesn't seem to you that the clerk is kind. I wonder, however, if a case of this sort is psychologically possible and, if it is, whether the belief that the clerk is kind shouldn't be taken to be a *pathological* and so *irrational* state.

Siegel doesn't accept (CT) presumably because she sees that its implementation would disqualify cognitively penetrated perceptual appearances from counting as conclusions of inferences. Consider her account of Jill's fear-penetrated appearance that (*Q*) Jack is angry. Jill is said to *infer* it from her appearance that (*P*) Jack has a neutral face. But it is implausible that Jill could have the appearance that *Q because* she (mistakenly) appreciates that *P* is evidence for *Q*. It is strongly intuitive, instead, that in the scenario described by Siegel, the psychological processes yielding Jill's appearance that *Q* are *secluded* from whatever Jill might concurrently appreciate about *P* and *Q*. As Lyons (2016) observes, the processes that produce perceptual experiences are

> fast, automatic, determined by domain specific principles proprietary to the perceptual system in question, *relatively* immune to cognitive penetration from occurrent beliefs and

25 Cf. Boghossian (forthcoming).

> goals and the like, largely unconscious in [their] inner workings, and performed by systems that are relatively independent and isolable from other cognitive mechanisms. (16)

So whether Jill appreciates that *P* supports *Q* or has no attitude about this issue, these facts would have no impact on the processes in charge of the production of Jill's perceptual appearance that *Q*.

These considerations deliver an even more devastating blow to Siegel's account. Siegel insists that cognitively penetrated appearances result from "inferences" that *the person* carries out 'covertly, silently, and unreflectively' (2017: 17). Yet Lyons' description strongly suggests that these transitions, if existing at all, must take place at *subpersonal* level. The subject as such doesn't perform them — rather, it is some perceptual system of the subject that makes these transitions for her.[26] Given this, it is very hard to understand how *the subject* could modulate the justifying power of her seemings by drawing relevant "inferences", along the lines outlined in Siegel (2017). It is utterly plausible that the subject cannot do it. Because of this, Siegel's account of the epistemic consequences of cognitive penetration appears seriously flawed.

Let me make a final objection to Siegel's account. Siegel claims that whether or not a cognitively penetrated seeming that *P* of *S* prima facie justifies *S*'s believing *P* depends on whether the seeming has positive or negative epistemic charge. Saying that a seeming of *S* has positive or negative epistemic charge is substantially the same as saying — for Siegel — that *S* holds the seeming with a high or low degree of rationally, as the conclusion of a good or bad mental transition (cf. 2017: §2). Siegel is thus committed to maintaining that cognitively penetrated appearances can be held by *S* (more or less) *rationally*. Since the type of rationality Siegel speaks about is *epistemic*, and epistemic rationality is responsive to *defeating evidence*, Siegel is also committed to the claim that any cognitively penetrated appearance that *P* held by *S* rationally is

26 I follow Lyons (2016) and Tucker (forthcoming) in interpreting 'personal' as 'attributable to the subject' and 'subpersonal' as, roughly, 'attributable only to parts of the whole subject'.

responsive to defeating evidence. This means that *S* must in principle be able to *revise* her appearance in light of any evidence *E* that discredited *P*'s accuracy or the reliability of the process that has produced the appearance. (Similarly, if a *belief* that *P* is held by *S* rationally, *S* should in principle be capable of revising it in light of any defeating evidence.) However, it is very dubious that *S* could revise — i.e. change or stop having — an appearance that *P* in light of *E*.[27] As observed in §2.3, our seemings typically don't change even when we become aware that they are deceptive or unreliable (cf. Brogaard 2013), and it would be only capricious to insist that *cognitively penetrated* seemings behave differently.[28] Since *S* could not revise any cognitively penetrated appearance in light of defeating evidence, *S* could not hold any such appearance rationally — *S* could entertain cognitively penetrated appearances only *arationally*.[29] This flies in the face of Siegel's contention that cognitively penetrated appearances can be held rationally.

Siegel (2017: 34-35) is aware that her account lays itself open to criticism of this type and attempts a response. She reminds us that *delusional* beliefs — such as the ones in Capgras syndrome or in schizophrenia — cannot be revised by the subject in light of defeating evidence. For these pathologies prevent the subject from doing it. Siegel stresses that these delusional beliefs nonetheless appear to be paradigms of *irrationality* — namely, *poor rationality* — rather than *arationality*. She infers from this that epistemic rationality doesn't require revisability.

I'm not convinced by this response. The intuition that *S*'s holding a belief with some degree of rationally requires *S*'s capability of revising it in the presence of defeaters is very widely shared in epistemology. If a plausible interpretation of Siegel's apparent counterexample compatible with this entrenched intuition were available, we should embrace it. Let me offer such an interpretation.

---

27 *S* could certainly look away or cover her eyes. But the resulting perceptual change wouldn't intuitively count as a *rational* revision (cf. Siegel 2017: 34).

28 One might suspect that if they are revisable, they aren't proper seemings but *doxastic* states.

29 Siegel insists that although *S* couldn't make herself stop having a given perceptual appearance in light of defeating evidence, *S* can normally "revise" the appearance in the sense of ceasing to rely on it in reasoning and action. Siegel also suggests that through perceptual learning, *S* might arrive at controlling the conditions that tend to give rise to a given type of appearance (cf. 2017: 35-37). All this looks plausible. However, it doesn't seem to me that these forms "revisability" and control are good diagnostics of the fact that the relevant appearances are held by *S* rationally.

Siegel would seem to confuse the notion of rationality as *epistemic rationality* — the one at stake here — with the notion of rationality as *coherence* between propositional attitudes (as characterized above in §2.4). In accordance with the second notion, a subject *S* can be said to be irrational for having two incoherent propositional attitudes simultaneously — e.g. a belief that *P* and an appearance that ¬*P*. *S*'s being irrational in this sense doesn't seem to require *S*'s being able to revise her incoherent attitudes. I suggest that people affected by unrevisable delusional beliefs are irrational just because *they have incoherent propositional attitudes* (e.g. a belief that *P* and evidence of some type that ¬*P)*. But they are not irrational in the sense of holding their delusional beliefs with a low degree of *epistemic* rationality. When we focus on epistemic rationality, we should probably conclude that these subjects entertain their delusional beliefs *arationally*, since they cannot revise them in light of recalcitrant evidence.

Evidence inferentialism

McGrath's and Markie's versions of evidence inferentialism explain the epistemic downgrade of badly cognitively penetrated appearances specifically in terms of the *basing* relation. These authors think that an appearance can be based on another appearance or experience in the same way as a belief can be based on another belief. The downgrade would happen when the content of the experience on which the appearance is based isn't adequate evidence for the content of the appearance.

McGrath (2013a, 2013b) suggests that there are two types of perceptual seemings: *receptive*, which represent only low-level properties, and *non-receptive*, which represent higher-level properties (cf. 2013b). McGrath submits that we produce non-receptive seemings from receptive seemings through a sort of inferential process. Receptive seemings are the inputs and non-receptive seemings are the outputs of *quasi-inferences*. According to McGrath,

(QI) A transition from a seeming that *P* to a seeming that *Q* is "quasi-inferential" just in case the transition that would result from replacing these seemings with corresponding beliefs that *P* and *Q* would count as genuine inference by the person.[30] (2013a: 237)

A transition counts as an inference by the person — for McGrath — only if (i) its input and output states are mental states of the person (rather than a sub-personal system) and (ii) there is an explanation 'that allows us to see the transition as the person's treating the content of the input state as supporting the content of the output state' (2013a: 238).[31] (ii) very strongly suggests that McGrath accepts that inferences should comply with (CT).

McGrath thinks that the tenet of phenomenal conservatism (PC) is true of all receptive seemings — hence, any receptive seeming that *P* of *S* provides *S* with prima facie justification for believing *P*.[32] Yet he submits that (PC) is false when applied to non-receptive seemings. The reason being that while receptive seemings are *given* to us, non-receptive seemings are *produced by us* through quasi-inferences. Because of this, McGrath contends — following Siegel — that non-receptive seemings can give us justification for believing their contents only if the relevant quasi-inferences that we perform are good.

Good and bad quasi-inferences can be characterized by a comparison with good and bad inferences. McGrath's grounding intuition is that a good quasi-inference from a seeming that *P* to a seeming that *Q* transmits *justifying power* from the first to the second seeming in the same way as a good inference from a belief that *P* to a belief that *Q* transmits justifying power from the first to the second belief. Justifying power transmits — whether we are speaking of seemings or beliefs — just in case the first state has justifying power and *P* sufficiently supports *Q*. McGrath suggests that all

30 For McGrath quasi-inferences also include mere dependence relations between seemings.
31 McGrath suggests such an explanation could appeal, for instance, to the person's grasp (good or faulty) of the support that one proposition gives to the other, the person's background information, or the cognitive states that make the person jump to conclusions.
32 I'm assuming that *S*'s seeming that *P* is clear and firm.

poor quasi-inferences are cases of jumping to conclusions caused by some interfering factor. They are cases in which *S* quasi-infers a seeming that *Q* from a seeming that *P*, where *P doesn't* support *Q*, because *S* is misguided by other mental states — e.g. fears, desires, expectations, etc. (cf. 2013b).[33]

McGrath explains our intuitions about the epistemic bearing of bad cognitive penetration by appealing to quasi-inference: the epistemic downgrade of a cognitively penetrated seeming would happen when the quasi-inference that has produced it is poor. Take *Angry Jack* for example. McGrath suggests the following rendition: Jill has a receptive seeming that (*P*) Jack's face displays certain *neutral* features. Nevertheless, under the influence of her unjustified belief that Jack is angry, Jill quasi-infers from the first appearance a non-receptive seeming that (*Q*) Jack is angry. This is a bad quasi-inference because *P* isn't evidence that *Q*. So Jill's seeming that *Q* cannot justify Jill's belief that *Q*.

McGrath's framework could be adduced to explain why good and bad cognitive penetration seem to have divergent epistemic consequences. Take again *Prospectors.* Gus is the expert and Virgil the novice. McGrath could argue that Gus' belief that

(*Q*) this pebble is gold

is prima facie justified by Gus' non-receptive seeming that *Q* because the latter seeming results from a *good* quasi inference. Specifically, Gus would entertain a receptive seeming that (*P*) this pebble has features *Fs*. From this, Gus would quasi-infer the non-receptive seeming that *Q*. Since *P in conjunction with Gus' background information* supports *Q*, this is a good quasi-inference. What said doesn't apply to Virgil, who doesn't possess Gus' background information. Virgil's quasi-inference from his receptive seeming that *P* to his non-receptive seeming that *Q* would thus be a

33 McGrath (2013a) says that these are cases of *free enrichment* — i.e. cases in which a non-receptive seeming is *freely* enriched due to an interfering state.

bad one. For *P alone* isn't evidence that *Q*. This would explain why Virgil's belief that *Q* isn't justified by his seeming that *Q*.

McGrath's evidence inferentialism can be seen as a special version of Siegel's process inferentialism, for it concentrates on *one* type of "inferences" (the appearance-to-appearance ones) and *one* way in which these transitions can fail (i.e. jumping to conclusions). Given its limited ambitions, McGrath's account isn't exposed to all the problems that afflict or may afflict Siegel's. For instance, McGrath doesn't claim that quasi-inferences are *genuine* inferences. Nor does he endorse controversial theses such as that there are unconscious experiences or that fears and desires are proper inputs to inference.

McGrath's account is nevertheless problematic in various respects. For example, note that quasi-inferences involve two states (i.e. two seemings) *both* with propositional content. Cognitive penetration might nevertheless depend on a transition from a state *without propositional content* to a seeming. Suppose for instance that, in *Angry Jack*, Jill has no receptive seeming about Jack's face. She has some *raw visual impressions* lacking conceptual content which, together with her unjustified belief that (*Q*) Jack is angry, brings about her seeming that *Q* (cf. Lyons 2016). McGrath's inferentialism couldn't explain cases of this type. Another limitation of it hinges on the possibility that cognitive penetration could *directly* affect receptive seemings (or raw sensations). Suppose that, because of her unjustified belief that *Q*, Jill has a *receptive* seeming that (*R*) Jack's face has *anger* features. If this is what actually happens in *Angry Jack*, McGrath should conclude that Jill is justified in believing that *Q* on the basis of her non-receptive seeming that *Q*. For this non-receptive *is* supported by the receptive seeming that *R* (cf. Lyons 2016). Some would find this counterintuitive.

McGrath's evidence inferentialism is afflicted by more serious difficulties. McGrath is convinced — correctly, in my view — that if a person *S* performs an inference, *S* must treat the content of the input state as supporting the content of the output state. This presumably means that

S must produce the output state because she appreciates that its content is supported by that of the input state. So McGrath endorses (CT) as a necessary condition for inference. Given this, the application range of McGrath's framework cannot but be drastically limited. For only a very few seeming-to-seeming transitions, if any, qualify as quasi-inferences on (QI). Let me explain why.

Consider that if *S* has been subjected to perceptual learning, the fact that *S* has a seeming that *Q* on the basis of a seeming that *P might* partly depend on her possessing a *background belief* that

(RI) *P*'s truth is a reliable indicator of *Q*'s truth.

Imagine for instance that *P* and *Q* respectively describe low-level and higher-level properties of birds, and that *S* has acquired the background belief (RI) in a birdwatching class before doing any fieldwork. Suppose that this background belief together with some fieldwork has eventually endowed *S* with a disposition to have a seeming that *Q* upon having a seeming that *P*. A seeming that *Q* acquired by *S* thanks to this disposition would be cognitively penetrated by *S*'s background belief (RI). Take now a transition from a seeming that *P* to a seeming that *Q* of this type in *S*'s mind. One could insist that if the seeming that *P* and the seeming that *Q* were replaced respectively by a belief that *P* and a belief that *Q*, the resulting belief-to-belief transition in *S*'s mind would satisfy (CT). For *S* would believe *Q,* via this transition, in virtue of her believing (RI)*,* and so in virtue of taking *P* to support *Q*. If this conclusion is correct,[34] McGrath's framework might account for cases of cognitive penetration of this type.

However, note that background beliefs of the sort just described are seldom possessed even by persons who have undergone various types of perceptual learning. We very rarely have beliefs about how low-level properties and higher level properties represented in our experiences are linked with one another, although a large amount of information of this sort is *implicitly* and *unconsciously*

34 This might perhaps be questioned. One might rejoin that in these circumstances, though *S* would actually take *P* to support *Q,* it is unclear that *S* would come to believe *Q* in virtue of her taking *P* to support *Q*. For it is still unclear that in these circumstances *S* would come to believe *Q* through her *agency* rather than a mere *subpersonal mechanism of association* partly shaped by (RI).

stored in our subpersonal perceptual systems (cf. Lyons 2016). In all cases of cognitive penetration in which the relevant seeming-to-seeming transition doesn't depend on some background belief of the type considered, if the seemings were replaced with corresponding beliefs, the resulting belief-to-belief transition would *not* satisfy (CT). Consider in fact that — as emphasized before — any transition from a seeming that *M* to a seeming that *N* in *S*'s mind would take place independently of any concurrent attitude of *S* about the relation of support between *M* and *N*. Thus, if the seemings were replaced with beliefs, the resulting belief-to-belief transition would also take place independently of any concurrent attitude of *S* about the relation of support between *M* and *N*. Hence, (CT) wouldn't be satisfied. These seeming-to-seeming transitions wouldn't qualify as quasi-inferences on (QI). McGrath's framework appears unable to account for all cases of cognitive penetration of this type, which are the largest majority.

Here is a last, probably lethal difficulty. McGrath is persuaded that since (QI) models quasi-inferences on proper inferences (between beliefs), and the latter are transitions *by the person*, quasi-inferences are also transitions by the person (cf. 2013a: 238). McGrath is indeed committed to the last claim because his explanation of the justifying power of non-receptive seemings rests — as we have seen — on the assumption that these seemings are produced *by us*. But this appears false. As Lyons (2016) has suggested, it is plausible that all transitions leading to the production of perceptual seemings happen at *subpersonal* level. Consider for example Gus in *Prospectors*. Despite Gus possesses accurate background information about how gold looks like and is able to detect gold successfully, it would be very odd to say that *it is Gus* — rather than his visual cognitive system — that works out an *appearance* that this nugget is gold from his *appearance* that this nugget has such and such features. (Although it would be fully appropriate to say that it is Gus who works out a *belief* that this nugget is gold from his *belief* that this nugget has such and such features.) If Lyons is

right — and I think he is — quasi-inferences like those hypothesized by McGrath, which are attributable to the subject's doing, simply don't exist.[35]

Let's turn to Markie's version of evidence inferentialism. Markie (2013) presupposes a non-unitarian view of perceptual appearance according to which perceptual appearance is a compound of sensations and seemings. To develop his account of cognitive penetrability, Markie mainly analyses *Prospector*. The impression that some have is that Gus' seeming that

(*Q*) this nugget is gold

prima facie justifies Gus' belief that *Q,* whereas Virgil's (assumed) identical seeming that *Q* doesn't even prima facie justify Virgil's belief that *Q*. Markie thinks that the best way to explain this impression is the following: Gus does know how to visually identify gold, and his seeming that *P* is just an instance of this knowledge-how. That's why Gus' seeming does justify the corresponding belief. Contrarily, Virgil doesn't know how to visually identify gold. So Virgil's seeming that *Q* cannot be an instance of any such knowledge-how. That's why Virgil's seeming doesn't even prima facie justify the corresponding belief.

This diagnosis leads Markie to reject (PC) and endorse a variant according to which:

(PC*) If *S* has an appearance that *P* that constitutes an instance of *S*'s knowledge-how to perceptually ascertain that *P, S* has thereby prima facie justification for believing *P*. (Cf. 2013: 250 and 262).

Markie calls *epistemically appropriate* the perceptual appearances that satisfy the left-hand side of (PC*) and so have justifying power. He elucidates the notion of knowledge-how to perceptually ascertain that *P*[36] as follows:

35 For further criticism see Lyons (2016), Huemer (2013a) and Siegel (2013c).
36 Markie (2013: 264) in fact focuses on the notion of knowledge-how to perceptually identifying something as being *Q*. The property I'm considering is an innocuous generalization of Markie's.

(KH) *S* knows how to perceptually ascertain that *P* if *S* has a disposition to entertain an appearance that *P* in response to *S*'s attending to certain features *Fs* of her experience, and *S* has this disposition in virtue of her having justified background information that these *Fs* indicate that *P*. (Cf. 263-264)

A few remarks are in order. According to Markie, *S*'s knowledge-how to perceptually ascertain that *P* need not come with *S*'s *reliable* practice. (In the Matrix, Gus would still know how to ascertain that a pebble is gold, though he would fail to ascertain it reliably.) Furthermore, *S*'s background information that *Fs* indicate that *P* need not consist of *S*'s *conscious beliefs*. *S* might just have evidence (e.g. experiences, justified beliefs) that provides *S* with prima facie justification *for believing* that if *Fs* are instantiated, then *P*. Finally, saying that *S* has the mentioned disposition in virtue of having this background information is not saying that this information *causes S* to have the disposition. It is only saying that the information helps to determine the *character* of the disposition and to sustain it.

Suppose now *S* has a disposition to entertain an appearance that *P* in response to her attending to features *Fs* of her experience, and she has this disposition in virtue of having justified background information that these *Fs* indicate that *P*. Thanks to (KH), any so-generated appearance that *P* would count as an instance of *S*'s knowledge-how to perceptually ascertain that *P*. Furthermore, thanks to (PC*), any such appearance would be epistemically appropriate, to the effect that it would give *S* prima facie justification for believing *P*.

We have seen how this framework applies to *Prospectors*. Markie suggests that it can be used to illuminate why Jill's seeming that (*P*) Jack's is angry, in *Angry Jack*, cannot justify Jill's belief that *P*. In Markie's interpretation, when Jill examines Jack's face, she attends to certain features *Rs*. In the envisaged scenario, Jill has *no* background information that these *Rs* indicate that

the subject — Jack, in the case in point — is angry. So Jill has no background information that these *Rs* indicate that *P*. However, Jill's irrational belief that *P* causes her having a seeming that *P*. This seeming isn't epistemically appropriate because it is not an instance of *S*'s knowledge-how to perceptually ascertain that *P.* Thus, this seeming cannot justify *S*'s belief that *P.*

We saw that a possible drawback of McGrath's evidential inferentialism is that it wouldn't explain cognitive penetration depending on a transition from a state *without propositional content* to a seeming. Markie's version of evidential inferentialism isn't subject to this problem, for in Markie's account, *S*'s seemings are thought to be based on experiences of *S* that don't need to have propositional content. Yet Markie's account is afflicted by another difficulty that troubles McGrath's evidential inferentialism too: cognitive penetration could *directly* affects the features *Fs* of experience that *S* attends to, rather than the correlated seeming. If this happened, *S*'s seeming would count as epistemically appropriate, though it might appear intuitive that it isn't so. (Suppose Jill attends to *anger* features, directly caused by her irrational belief.)[37]

Markie's account isn't hostage to problems related to (TC), as it doesn't appeal to inferential transitions between beliefs. A serious difficulty of Markie's account is, however, that what Markie (2013) considers to be knowledge-how doesn't actually appear to be knowledge-how. Specifically, (KH) takes *S*'s having a disposition to entertain an appearance that *P* in given circumstances to be a condition sufficient for *S*'s having knowledge-how to perceptually ascertain that *P.* This principle is very dubious. The problem is that knowledge-how is customarily conceived of as a type of goal-directed ability possessed by *agent* (see for instance Carr 1998, Markie 2006 and Cath forthcoming), but *S*'s disposition is not a goal-directed ability possessed by *S* as an agent. *S*'s disposition appears to be a mere "mechanistic" or "automatic" function of *S*'s perceptual cognitive system. Because of this problem, Markie's account of the epistemic effects of cognitive penetration proves very implausible. Let me expand on this.

37 Markie (2013: 266) bites the bullet and insists that this upshot is perfectly acceptable.

Whenever *S* knows how to do something *G* (e.g. riding a bike, counting, walking, etc.), *S* is capable of adopting a behaviour directed to achieve the goal *G*. Accordingly, *S* must have a correct understanding of *G* and be able to follow the norms that allow her to pursue it. The most basic of these norms have presumably this form: to attain *G*, in circumstances *C*, do *D*. For instance, if *S* knows how to ride a bicycle, *S* must be able to behave in a way directed to the goal of riding a bicycle. *S* must thus correctly understand this goal and be able to conform to the correlated norms. One of these norms — let's call it *N* — might state this: in order to ride a bicycle, if you lose your balance while cycling, regain your balance by leaning in the opposite direction. Note that *S*'s following *N* doesn't require *S* to make an explicit reasoning that appeals to *N* to decide what to do (cf. Markie 2006). The only thing required of *S* in order to follow *N* is presumably this: when *S* finds herself off balance while cycling, as long as *S* intends to ride the bicycle, *S* should somehow *see* or *feel*[38] that she needs to lean in the opposite direction to regain balance, and she should act accordingly. (This seems to apply, mutatis mutandis, to any norm of knowledge-how in general.)

When *S* exercises her knowing how to do something *G*, if *S* were asked why she has behaved in a given way, in some cases *S* could correctly explain, upon reflection, why she has behaved that way by giving an approximate description of the norms she has followed. For instance, if asked why she has leant left while riding her bike, *S* might say that she has done so because she needed to regain balance (cf. Markie 2006). In other cases, *S*'s vocabulary may not be rich enough to describe the norms *S* has actually followed. (Imagine the question asked to *S* concerns her knowing how to *whistle* a melody.) Also, *S* might lack reflexive or linguistic abilities necessary to supply explanations altogether. (Imagine *S* is a small child who knows how to clap.)[39]

38 Cath (2012) for instance suggests that *S* should have a *seeming* that she ought to act in a certain way.
39 Furthermore, the norms reported by S might prove uninformative, for they might just refer to inexpressible feelings. (Suppose the question asked to S concerns her knowing how to ascertain that a number of objects on the table is higher than 10 without counting them.) And even if S didn't lack these abilities, she might make mistakes of various type in reporting her internal states.

Suppose now that *S* has an appearance that *P* as a result of her exercising a disposition of the type described in (HK). Markie would maintain that this appearance is an instance of *S*'s *knowing how* to perceptually ascertain that *P*. This seems false. It is very plausible that if *S* were asked why she has brought about the appearance that *P* upon attending to features *Fs* of her experience, *S* could offer no description of the norms she has followed. This wouldn't depend on *S*'s inadequate vocabulary or lack of introspection. The reason would simply be that the seeming that *P* would pop up in *S*'s mind *mechanically*, without *S* following any norm to generate it (not even in the undemanding sense described before). The production of this seeming would not be accompanied in *S*'s mind by the characteristic features of goal-directed acts. All this strongly suggests that the production of this appearance would not involve *S as an agent.* Since *S*'s seeming that *P* would not result from an agential ability of *S,* it would be inappropriate to take it to be an instance knowledge-how of *S* (cf. Lyons 2016).[40]

In conclusion, the central problem of Markie's inferentialism is the same as the central problem of McGrath's and Siegel's versions of it. These three authors are convinced that whether or not a cognitively penetrated appearance has justifying power essentially depends on whether and how the subject, conceived of as an agent or person, has brought about the appearance. This thesis is very implausible, for appearances in general are not brought about by persons or agents. Markie, McGrath and Siegel have supplied no good reason to believe the contrary. Since this crucial presupposition of the internalist accounts of the epistemic import of cognitive penetration defended by inferentialists is extremely implausible, internalists had better look for an alternative account.

3.6 Taming cognitive penetrability

Consider any perceptual appearance that *P* of *S* cognitively penetrated such that *S* couldn't detect it by simply inspecting the appearance. (I expand on this notion below.) Phenomenal conservatives

40 For further criticism see Ghijsen (2016) and Lyons (2016).

can contend that any such perceptual appearance retains the power to prima facie justify *S*'s believing *P*. This contention should go hand in hand with *explaining away* the intuitions and impressions incompatible with it that many epistemologists claim to have.[41] Phenomenal conservatives have all the trumps to pursue both tasks successfully. Let's start with the first.

As we have seen, a perceptual seeming that *P* of *S* might happen to be cognitively penetrated by a state of *S* because this state might directly causes in *S* a *high-level* representation that *P*. In these situations, there might be a mismatch detectable by *S* between this high-level representation and other lower-level representations component parts of the same perceptual seeming. Sometimes *S* might be actually aware of the mismatch. (Suppose *S* is a microscopist who endorses preformationist and observes a sperm cell. *S* might actually realize that although she doesn't perceive certain parts of the cell as components of an embryo, she sees their sum as an embryo. Another case might concern a variant of *Angry Jack*. Jill might actually be aware that although she cannot spot any element expressing anger in Jack's face, she sees the face as angry.) In other cases it might happen that *S* isn't aware of any incoherence within her cognitively penetrated seeming, though it would be easy for her to spot a discrepancy if she only attended to some of its features. In both these general cases, *S*'s perceptual appearances would be cognitively penetrated in a way that *S* could detect an effect of it by simply inspecting the appearances. Let's say that in these cases *S*'s perceptual seemings would be *detectably* cognitively penetrated.

In other cases of cognitive penetration, nevertheless, it might be difficult if not impossible for *S* to find any incoherence in her perceptual seemings. These cases encompass the remaining ones in which cognitive penetration affects only *S*'s high-level perceptual contents, and all those in which cognitive penetration affects *S*'s perceptual seemings as wholes. In all these cases *S*'s perceptual seemings would be cognitively penetrated so that *S* couldn't detect any effect of it by

41 It is important to stress that certain epistemologists — e.g. Lycan (2013) and Huemer (2013a) — don't share some of these intuitions.

simply inspecting the appearances. Let's say that in these cases *S*'s perceptual seemings would be *undetectably* cognitively penetrated.

Detectable cognitive penetration affects negatively perceptual justification but it doesn't seem to raise particular difficulties to phenomenal conservatism. The phenomenal conservative can acknowledge that if *S*'s perceptual seeming that *P* were detectably cognitively penetrated, *S*'s belief that *P* wouldn't be perceptually justified. Huemer (2013a, and more explicitly 2013b: 345) suggests that, in this case, *S*'s justification for *P* based on *S*'s high-level content that *P* would be *defeated* by some concurrent lower-level content of *S*'s experience that doesn't support *P*. I suggest an alternative diagnosis: *S*'s *incoherent* perceptual seeming that *P* wouldn't provide *S* with even *prima facie* justification for believing *P*. In fact recall that on (PC) — as characterized in §2.2 — only a *clear and firm* seeming that *P* of *S* could give *S* prima facie justification for believing *P*.

Let's turn to *undetectable* cognitive penetration. Hereafter, whenever I use the expression 'cognitively penetrated seeming(s)' I always refer to *undetectably* cognitively penetrated seeming(s). The best argument I can think of to conclude that any cognitively penetrated seeming supplies the subject with prima facie justification for believing its content draws from McGrath's case for (PC) detailed in §2.4. It appears very plausible that if *S* had a cognitively penetrated seeming that *P* and no defeater, the only *epistemically rational* doxastic attitude from *S*'s own standpoint that *S* could have towards *P* is *believing P*. If we rely on the internalist intuition that *S*'s believing *P* is *epistemically rational* only if it is *epistemically justified*, we must conclude that in this case *S* has justification for believing *P.* Precisely, we must conclude that if *S* has a cognitively penetrated appearance that *P* and no defeating evidence, *S* has both prima facie and ultima facie justification for believing *P.* Thus, *S* has prima facie justification for believing *P* (cf. Huemer 2013a, 2013b).

Here is a concrete application of this reasoning, described in Huemer (2013b). Suppose that a fearful subject *S* suspects that (*G*) there is a gun in the refrigerator. In fact the fridge contains only

a banana. However, when she goes to check, *S*'s fear causes her to see the banana as a gun. Imagine that *S* has a cognitively penetrated seeming that *G* and no reason to suspect that the seeming is inaccurate or unreliable. It is clear that *S* would be justified in believing *G*. As Huemer emphasizes:

> This does not strike me as a difficult or borderline case; it strikes me as a perfectly clear case of epistemic justification. If things look that way to you, and you have no reason to doubt your eyes, then you would be crazy not to think there is a gun in the refrigerator. (743).

To lend further support to the thesis that our cognitively penetrated appearances give us prima facie justification for believing their contents, Huemer (2013b) adduces another thought experiment. The experiment aims to exemplify the following argument: since a subject *S* wouldn't be able to find any epistemically significant difference between a cognitively penetrated and a non-cognitively penetrated[42] appearance if *S* inspected both of them, the appearances of these two types cannot differ in terms of their intrinsic justifying power. Hence, both of them must provide *S* with prima facie justification for believing their contents.

Let *G* be 'there is a gun in the fridge' and let *E* be 'there is a carton eggs in the fridge'. In this new scenario, *S* has a perceptual appearance that *G and E*. Suppose that the part of *S*'s appearance that represents *G* is cognitively penetrated by *S*'s fear of guns (the gun is actually a banana), whereas the part that represents *E* isn't cognitively penetrated. Imagine that after scrutinizing her appearance and considering all information available to her, *S* is unable to find any reason to distrust *any* part of the appearance.[43] As a result, *S* comes to believe *E.* In this case, *S*'s believing *E* appears to be rational — at least prima facie so. However, suppose that *S* responds to

42 Neither detectably nor undetectably.
43 Suppose *S* doesn't suspect that her fear could penetrate her experience.

her perceptual appearance and inability to find reasons to distrust it by *refusing* to believe *G* — i.e. by disbelieving *G* or suspending judgment about it. This attitude of *S* towards *G* is absurd. For since '*S* would have no rational way of explaining why *S* believed *E* while refusing to [believe] *G*, then *S* [is] irrational to believe *E* while refusing to [believe] *G*' (746). Clearly, in these circumstances, *S*'s believing *G* would be rational — at least prima facie — just as *S*'s believing *E* is. This involves that *S*'s believing *G* would be prima facie justified, and it would be so because of *S*'s having the cognitively penetrated appearance (or appearance-part) that *G*.

While I find this thought experiment straightforward, Siegel (2013c) finds it is questionable. She thinks that Huemer in his reasoning makes use of a principle like this:

(R1) A doxastic attitude of *S* is rational only if *S* has a rational way to explain why she adopts it. (Cf. 753)

Siegel claims that (R1) is quite dubious. To clarify why, she appeals to an example originally due to McGrath (2013b). Imagine that *S* has an occurrent apparent memory that (*P*) her own child plays the piano better than any other child. Suppose that *S* also has an *unretrieved* memory that (*Q*) another child plays the piano better than her own child. If *S* refused to believe *P* in these circumstances, *S*'s attitude would be — according to Siegel — *rational.* For *S*'s memory that *Q* is — intuitively — a defeater of *S*'s support for *P* based on *S*'s apparent memory that *P*. Yet since *S*'s memory that *Q* is unretrieved*,* if *S* refused to believe *P, S* would have no rational way to explain her attitude. This clashes with (R1).[44]

An austere internalist accessibilist might respond that given that *S*'s memory that *Q* is actually unretrieved*,* it doesn't intuitively count as a defeater. But this response would certainly be

44 Siegel (2013c: 753) offers an additional example. My reply to the young pianist example could easily be adapted to respond to it.

controversial. What makes Siegel's objection flawed is that (R1) is a very general proposition that Huemer is not committed to. Huemer's thought experiment can legitimately be interpreted as relying on, not (R1), but this more specific principle:

(R2) A doxastic attitude *A* of *S* is prima facie rational in virtue of an occurrent experience *E* of *S* only if *S* could rationally explain why she adopts *A* by appealing to *E*.

(R2) looks straightforward. In the young pianist example, if *S* refused to believe *P,* this attitude would *not* be prima facie rational in virtue of *S*'s apparent memory that *P*.[45] This is attested by (R2): in the envisaged situation, *S* couldn't rationally explain why she would refuse to believe *P* by appealing to her apparent memory that *P*. On the other hand, if *S* believed *P,* this attitude would intuitively be prima facie rational in virtue of *S*'s apparent memory that *P*. Note that in this case, in harmony with (R2), *S* could rationally explain why she would do so by adducing her apparent memory that *P*.

Importantly, once (R1) is replaced with (R2), Huemer's thought experiment looks impeccable. In the situation imagined by Huemer, if *S* refused to believe *G,* this attitude would not be prima facie rational in virtue of the part *G* of her appearance. This is certified by (R2): *S* couldn't rationally explain why she would refuse to believe *G* by appealing to the part *G* of her appearance. Contrarily, if *S* believed *G*, this attitude of *S* would intuitively be prima facie rational in virtue of the part *G* of *S*'s appearance. In this case, *S* could rationally explain why she would do so by appealing to the part *G* of her appearance, in accordance with (R2).

To bring to completion the task of defending (PC) from cognitive penetration objections, the phenomenal conservative can concede that there is *something* wrong with beliefs based on appearances that are (undetectably) badly cognitively penetrated. Beliefs like these look in some

45 I assume that this apparent memory is an experience.

sense epistemically defective. The phenomenal conservative should insist, however, that it is quite another issue as to whether these defects concern *prima facie justification*. More explicitly, the phenomenal conservative should contend that the intuitions that seem to suggest that beliefs of this type lack prima facie justification are misconstrued, for these beliefs actually lack *a different epistemic property* that can be confused with prima facie justification.

To start with, phenomenal conservatives might be allured by the view that in every case of bad cognitive penetration, *S*'s seeming that *P* doesn't supply *S* with *all things considered* justification for believing *P* because some mental state of *S* is a *defeater* of *S*'s prima facie justification. In accordance with this view, the cases of bad cognitive penetration are situations in which *S* has some belief, memory, intuition or a similar state that counts for *S* as a reason to take her appearance to be untrustworthy. For example, in *Prospectors*, Virgil would probably be aware that he lacks experience and training. This would give him a reason to suspect that his perceptual appearance that the nugget is gold is untrustworthy (cf. Tucker 2010: 539). The fact that Virgil lacks ultima facie justification would explain why his belief that the pebble is gold looks epistemically defective, and why it appears to be epistemically worse than Gus' identical belief, which is all things considered justified (cf. Georgakakis and Moretti 2019).

Unfortunately, this account cannot generalize to all cases of bad cognitive penetration because there is actually no guarantee that *S* would always have a defeater of her seeming-based justification in these circumstances. This is evident as we consider cases different from *Prospectors*. Take *Angry Jack*. It is possible and not implausible that in a scenario of this type, Jill could be unaware that her appearance might be caused by some of her prior beliefs. It appears false, therefore, that Jill would inevitably have a reason to take her appearance to be untrustworthy in these circumstances.[46] Markie indicates a deeper problem of this account: it seems to locate Virgil's error in the wrong place. 'It is not that Virgil shouldn't form his belief on the basis of his seeming

46 See Siegel 2012 for further discussion.

experience, because its epistemic support is defeated. [The impression] is that he should not have his seeming experience in the first place' (2013: 258).

Tucker (2011) and Huemer (2013b) make a different proposal to the effect that whenever a seeming that *P* of *S* is badly cognitively penetrated, *S*'s belief that *P* based on it is prima facie justified but not *warranted*. Since warrant is the property that must be added to true belief to become knowledge, according to this proposal, whenever *S*'s seeming that *P* is badly cognitively penetrated, *S*'s belief that *P* based on it cannot result in *S*'s *knowledge* that *P*. A warranted belief is typically thought of as one that isn't just accidentally true, for it is *properly* connected with the external world by a reliable or truth-tracking belief production mechanism (cf. Plantinga 1993). It is intuitively plausible that neither Jill's perceptual belief in *Angry Jack* nor Virgil's perceptual belief in *Prospectors* is caused by a reliable belief-forming process. On the other hand, Gus' perceptual belief sustained by his expertise seems to be produced by a reliable belief-forming process.

Markie (2013) has worked out a variant of *Prospector* that casts doubts on this explanation. Suppose both Virgil and Gus are in the Matrix and are fed with sensory inputs identical to those that they would have in the real world. Gus' seeming that his pebble is gold is still partly caused by Gus' skills and background information, so it still appears to be affected by *good* cognitive penetration. On the other hand, Virgil's identical seeming is still partly caused by Virgil's desire to become rich, so it still appears to be afflicted by *bad* cognitive penetration. Hence, Virgil's seeming-based belief still looks epistemically worse than Gus' identical belief, despite *both* Virgil and Gus lack perceptual knowledge and warrant now.[47] For both Virgil's and Gus' beliefs are *disconnected* from the world, due to the sceptical scenario. This suggests that what explains the epistemic inadequacy of a belief based on a badly cognitively penetrated seeming isn't the fact that the subject lacks warrant for it.

47 Markie only mentions Virgil's and Gus' lack of knowledge, but he would agree that they both also lack warrant.

Pressed by Markie's objection, Tucker (2010, 2013) has offered another explanation. In Tucker's new account, when a seeming that *P* of *S* is badly cognitively penetrated, *S* possesses prima facie justification for believing *P.* However*, S* is *epistemically blameworthy* for having produced her seeming; thus *S* would also be *epistemically blameworthy* for believing *P* on the basis of it.[48] This proposal assimilates the situations in which *S* has a badly cognitively penetrated seeming to situations in which *S fabricates* her own evidence. This is a case of evidence fabrication imagined by Huemer (2013a: 344-345): suppose *S* has a brain-manipulation device invented by a scientist. Imagine that for some reason *S* intends to make herself believe that (*Q*) there is a cat before her. So *S* pushes a button on the device that causes *S* to have a hallucination that *Q* while simultaneously erasing *S*'s memory of her having used the brain-manipulation device. As a result, *S* believes *Q*.[49] In this example *S* appears to be epistemically blameworthy for producing her evidence that *Q*. *S* would also appear epistemically blameworthy for having a belief that *Q*, if she believed *Q* on the basis of her fabricated evidence. Still, *S*'s evidence provides *S* with prima facie (and all things considered) justification for believing *Q.* For, from *S*'s viewpoint at the time she has the hallucination, there is nothing about the hallucination that distinguishes it from a perfectly ordinary experience (cf. Tucker 2010, 2013 and Huemer 2013a). Tucker thinks that something very similar holds true in the cases of bad cognitive penetration: *S* is epistemically blameworthy for producing a cognitively penetrated appearance that *P,* and *S* would also be epistemically blameworthy for believing *P* if she did so on the basis of her appearance. Nevertheless, *S*'s belief would be prima facie justified by *S*'s appearance. All this would explain — according to Tucker — why Virgil's prima facie justified belief that *P* is intuitively epistemically worse than Gus' prima facie justified identical belief, even when both subjects are in the Matrix. Virgil 'is blameworthy for his false belief and his inappropriately-caused seeming and [Gus] isn't' (Tucker 2010: 541).

48 Tucker specifically focuses on *wishfully* produced appearances (like Virgil's appearance in *Prospectors*). I follow McGrath (2013b) in extending this proposal to cover virtually all cases of bad cognitive penetration.
49 See Tucker (2010) and McGrath (2013b) for alternative examples.

Markie (2013) complains that Tucker's new proposal leaves it unspecified the nature of the negative upshot for which *S* would be epistemically blameworthy when entertaining a badly cognitively penetrated appearance. In particular, Tucker doesn't clarify the sense in which a badly cognitively penetrated appearance would be *inappropriately* produced by *S*. I do have an independent concern: it is normally assumed that *S* can be blameworthy for doing something *X* only if *X* is a *voluntary* act of *S* or, at least, *S* has s*ome control* on *X*. Yet since the processes that make *S*'s seemings badly cognitively penetrated (supposing cognitive penetration exists) take place at *subpersonal* level*,* it is implausible that *S* could control them. *S* may be able to revise a *belief* of her formed on a seeming cognitively penetrated by her desires or expectations, if *S* discovered that the belief has these features. Yet it is quite dubious that *S* could *prevent* previous or concurrent desires or expectations of her from penetrating her seemings.[50] As we have already seen, it is equally dubious that *S* could *revise* her cognitively penetrated appearances. In conclusion, the claim that *S* would be blameworthy for entertaining a badly cognitively penetrated seeming appears odd and so implausible. This casts doubts on the acceptability of Tucker's proposal.

Let me outline an alternative account that aims to improve on Tucker (2011) and Huemer (2013b)'s proposal. Consider a belief that *P* of *S* based on a seeming that *P* of *S* that looks badly cognitively penetrated. Tucker and Huemer maintain that we have the impression that this belief is epistemically defective because we take it to be unwarranted. I suggest that we have this impression because — more precisely — we judge that the belief isn't produced by a *properly functioning* cognitive faculty of *S*. This is why we find the belief unwarranted. If this is correct, Markie's objection can be answered. For in the Matrix variant of *Prospectors*, only Virgil's belief is epistemically defective in this specific sense, not Gus'.

---

50 If perceptual learning actually produces good cognitive penetration, it might be correct to maintain that *S* can control factors that produce good cognitive penetration by controlling these learning process. Perhaps *S*'s appearances might also be subject to *bad* cognitive penetration depending on (bad) perceptual learning that *S* is responsible for. It might thus be correct to maintain that, *in these cases*, *S* is epistemically blameworthy for having the relevant appearances. However, note that these cases would be very different from the typical examples of bad cognitive penetration — such as *Angry Jack* and Virgil's case in *Prospectors* — discussed in the literature.

To appreciate how this account works, let me expand on the notion of proper function. Consider an example from Plantinga (1993a: 195-198 and 205-207). Imagine *S* has a brain lesion that elicits various cognitive processes generating random beliefs. One of these processes — call it *PR* — produces only the *true* belief that

(*Q*) One has a brain lesion.

Suppose that *PR* is *very reliable* because it would always result in one's having a true belief that *Q* if *PR* took place in one's brain. Despite *PR*'s high reliability, it is intuitive that *S*'s belief that *Q* cannot result in *S*'s *knowing* that *Q*. Plantinga suggests that we have this intuition because we realize that *S*'s belief that *Q* isn't formed through a *properly functioning* cognitive faculty. Plantinga concludes from this example that any reliabilist account of warrant must include a condition that states that a belief of a subject *S* can be warranted only if it is formed through cognitive faculties of *S* that function properly. Note that this is just a *necessary* condition. The proper functioning of *S*'s cognitive faculties doesn't suffice for warrant: it is possible that *S*'s faculties be functioning properly but *S*'s beliefs still lack warrant. This happens when *S*'s faculties and the environment in which *S* is are not properly attuned (cf. Plantinga 1993b: 7). Imagine for example that *S* is in a sceptical scenario. *S*'s cognitive faculties could in this case function properly, albeit *S*'s perceptual beliefs would remain unwarranted.

The notion of proper function presupposes the one of *design plan,* which specifies the way in which a thing is supposed to function in various circumstances to achieve certain goals. In the case in point a design plans specifies the ways in which *S*'s cognitive faculties are supposed to function in different circumstances to produce true beliefs. The existence of a design plan doesn't necessarily require the existence of a *conscious* designer, such as God (cf. 21). Millikan (1984), for instance, gives a fully naturalistic-evolutionary account of this notion. Plantinga (1993b: 237) ultimately endorses a theistic explanation.

Suppose now that *S* has a cognitively penetrated perceptual seeming that *P* and believes *P* on that seeming. I suggest that our intuitive judgments about the epistemic standing of *S*'s belief that *P* result from a process of this type: we (implicitly) attribute a design plan to the cognitive faculties of *S* involved in producing the belief that *P*. Then, we check whether the seeming that *P* has been formed through the proper functioning of those faculties. If the answer is positive, we conclude that the cognitive penetration of the seeming is good, and that *S*'s belief that *P* is not epistemically defective.[51] This is what presumably happens when we judge that Gus' belief in *Prospector* isn't epistemically defective. If the answer is instead negative, we conclude that the cognitive penetration of *S*'s seeming that *P* is bad, and *S*'s belief that *P* is epistemically defective as originated from faculties that aren't properly functioning. This is what presumably happens when we judge that Virgil's belief in *Prospector* and Jill's belief in *Angry Jack* are epistemically defective.

This account enables us to respond to Markie's criticism. When we scrutinize the Matrix variant of *Prospectors* described before, we acknowledge that both Virgil's and Gus' perceptual beliefs are unwarranted because they are secluded from the external world. Nevertheless, since we attribute to Virgil and Gus the same design plan that we would attribute to them if they were in the ordinary world, we do perceive Virgil's belief to be epistemically *worse* than Gus'. For we reckon that the belief isn't formed through *properly functioning* cognitive faculties, whereas Gus' belief is.

Markie (2013) thinks that an account of this type misses the target because:

> We can modify the [Matrix variant of *Prospectors*] so that Virgil meets the relevant external condition, keeping his internal mental state the same. Perhaps, his design plan calls for his desires to penetrate his perceptions... [But even so, Virgil] still forms his belief on the basis of a seeming experience resulting from the penetration of his visual perception by his desire.

51 As far as the bearing of cognitive penetration is concerned.

> That alone makes his belief epistemically inappropriate in a way in which Gus's belief is not. (259-260)

I don't find these sketchy considerations convincing. Let's try to produce a viable variant of *Prospectors* similar to the one that Markie might have in mind. In this variant, Gus is an expert gold prospector, whereas Virgil knows only a few basic features of gold. While Gus lives in the ordinary world, Virgil inhabits another world *W*. That world is almost identical to ours, with one difference: in *W* a strong desire that a pebble is gold is very likely to turn the pebble into gold. Suppose Virgil's cognitive faculties have been designed (by the nature or God) in a way to enable him to acquire many true beliefs. Given the mentioned peculiarity of *W*, Virgil's faculties are such that his strong desire that a pebble is gold can easily penetrate his perception of the pebble, to elicit in Virgil a (probably true) belief that the rock is gold. One day, Virgil and Gus are abducted and put in the Matrix by the crazy scientist. The scientist erases all traces of the abduction from their memories, so that the two prospectors keep thinking they are in their original worlds. In the virtual reality of the Matrix, Virgil and Gus come to inspect the same pebble and have an identical cognitively penetrated seemings that (*P*) the pebble is gold. Virgil' seeming is penetrated by his desire, while Gus' is penetrated by his background information and skills. Thereby, they both believe *P*.

In this thought experiment, the design plan of Virgil requires his desire that *P* to penetrate his appearance that *P*. Virgil's belief that *P* is thus produced by faculties that are *functioning properly*. Markie insists that, even in this case, Virgil's belief that *P* is intuitively epistemically worse than Gus' belief that *P*. I candidly confess that I lack the intuition that Markie seems to have, and I trust that many readers may lack it as well. In the envisaged scenario, Virgil and Gus acquire their respective perceptual beliefs that *P* while they are secluded from the external world. So both their beliefs are unwarranted and cannot constitute knowledge. Furthermore, both beliefs are produced by cognitive faculties that function properly, and they are both prima facie justified by

identical perceptual appearances. I don't see any clear sense in which one of these beliefs could be epistemically worse than the other.

3.7 Conclusions

I have scrutinized arguments against phenomenal conservatism, levelled by epistemic externalists and internalists, which appeal to the controversial thesis that the contents of perceptual appearances can be penetrated by previous or concurrent cognitive states of the subject. These objections adduce intuitions that seem to suggest that cognitively penetrated perceptual appearances often lack the ability to prima facie justify their contents. This would clash with (PC). In response, I have suggested that the adduced intuitions cannot be vindicated by the externalist contention that cognitive penetration often makes perceptual appearances unreliable. Furthermore, I have argued at length that the adduced intuitions cannot be vindicated by the internalist contention that cognitive penetration often makes perceptual appearances epistemically irrational or arational. I have also defended the claim that undetectable cognitively penetrated appearances don't lose their ability to prima facie justify their contents, and the claim that the intuitions that seem to suggest otherwise can be explained away as cases of misidentification of epistemic properties. In particular, I have suggested that the beliefs based on undetectable cognitively penetrated appearances are prima facie justified but they look defective when they are the product of cognitive faculties that appear not to function properly. In conclusion, cognitive penetrability doesn't seem to be a problem for phenomenal conservatism.

References

Bergmann, Michael. 2006. *Justification without awareness*. NY: Oxford University Press.

Boghossian, Paul. 2014. What is inference? *Philosophical Studies* 169: 1-18.

———— forthcoming. Inference, mechanism, and normativity. In *Reasoning: essays on theoretical and practical thinking*, eds. Magdalena Balcerak Jackson and Brendan Balcerak Jackson. Oxford: Oxford University Press.

Brogaard, Berit. 2013. Phenomenal seemings and sensible dogmatism. In *Seemings and justification: new essays on dogmatism and phenomenal conservatism*, ed. Chris Tucker, 270-289. New York: Oxford University Press.

Broome, John. 2014. Normativity in reasoning. *Pacific Philosophical Quarterly* 95: 622-633.

Carr, David, 1981. Knowledge in practice. *American Philosophical Quarterly* 18: 53-61.

Cath, Yury. 2012. Knowing how without knowing that. In *Knowing How: Essays on Knowledge, Mind, and Action*, eds. John Bengson and Mark A. Moffett, pp. 113–35. Oxford: Oxford University Press.

———— Forthcoming. Knowing how. *Analysis*.

Comesaña, Juan. 2002. The diagonal and the demon. *Philosophical Studies* 110: 249-266.

———— 2010. Reliabilist evidentialism. *Nous* 44: 571-600.

Firestone, Chaz and Scholl, Brian. 2016. Cognition does not affect perception: evaluating the evidence for 'top-down' effects. *Behavioral and Brain Science*s 39: 1-72.

Frege, Glottob. 1979. Logic. In *Posthumous writings* eds. Hans Hermes et al. Chicago: University of Chicago Press.

Fumerton, Richard. 2006. *Epistemology.* Malden: Blackwell Publishing.

Gatzia, Dimitria and Brogaard, Berit 2017. Pre-cueing, perceptual learning and cognitive penetration. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2017.00739.

Georgakakis, Christos and Luca Moretti. 2019. Cognitive penetrability of perception and epistemic justification. *Internet Encyclopedia of Philosophy*. https://www.iep.utm.edu/cog-pene/. Accessed 31 May 2019.

Goldman, Alvin. 1979. What is justified belief? In *Justification and knowledge*, ed. George Pappas, 1-25. Dordrecht: Reidel.

Goldman, Alvin and Beddor, Bob. 2016. Reliabilist epistemology. *Stanford Encyclopedia of Philosophy*, ed. Edward Zalta. https://plato.stanford.edu/archives/win2016/entries/reliabilism/.

Ghijsen, Harmen. 2016. The real epistemic problem of cognitive penetration. *Philosophical Studies* 173: 1457-1475.

Huemer, Michael. 2007. Compassionate phenomenal conservatism. *Philosophy and Phenomenological Research* 74: 30-55.

———— 2013a. Phenomenal conservatism Über Alles. In *Seemings and justification: new essays on dogmatism and phenomenal conservatism,* ed. Chris Tucker, 328-350. New York: Oxford University Press.

———— 2013b. Epistemological asymmetries between belief and experience. *Philosophical Studies*162: 741-748.

———— 2016. Inferential appearances. *In Intellectual assurance: essays on traditional epistemic internalism*, eds. Brett Coppenger and Michael Bergmann, 144-160. Oxford: Oxford University Press.

Johansson, Petter, Lars Hall, Sverker Sikström, Betty Tärning, and Andreas Lind. 2006. How something can be said about telling more than we can know: On choice blindness and introspection. Consciousness and Cognition 15: 673-692.

Locke, John. 1689/2008. *An essay concerning human understanding*. Oxford: Oxford University Press.

Long, Todd R. 2012. Mentalist evidentialism vindicated (and a super-blooper epistemic design problem for proper function justification). *Philosophical Studies* 157: 251-266.

Lycan, William G. 2013. Phenomenal conservatism and the principle of credulity. In *Seemings and justification: new essays on dogmatism and phenomenal conservatism,* ed. Chris Tucker, 293-305. New York: Oxford University Press.

Lyons, Jack. 2011. Circularity, reliability, and the cognitive penetrability of perception. *Philosophical Issues* 21: 289-311.

———— 2016. Inferentialism and cognitive penetration of perception. *Episteme* 13: 1-28.

Macpherson, Fiona. 2012. Cognitive penetration of colour experience: rethinking the issue in light of an indirect mechanism.' *Philosophy and Phenomenological Research* 84: 24-62.

Markie, Peter. 2006. Epistemically appropriate perceptual belief. *Nous* 40: 118-142.

———— 2013. Searching for true dogmatism. In *Seemings and justification: new essays on dogmatism and phenomenal conservatism,* ed. Chris Tucker, 248-268. New York: Oxford University Press.

McGrath, Matthew. 2013a. Phenomenal conservatism and cognitive penetration: the 'Bad Basis' counterexamples. In *Seemings and justification: new essays on dogmatism and phenomenal conservatism,* ed. Chris Tucker, 225-247. New York: Oxford University Press.

———— 2013b. Siegel and the impact for epistemological internalism. *Philosophical Studies* 162: 723-732.

McHugh, Connor, & Way, Jonathan 2016. Against the taking condition. *Philosophical Issues* 26: 314-331.

Millikan, Ruth. 1984. Naturalist reflections on knowledge. *Pacific Philosophical Quarterly* 4: 315-334.

Mole, Christopher. 2015. Attention and cognitive penetration. In *The cognitive penetrability of perception: new philosophical perspectives*, eds. John Ziembekis and Athanassios Raftopolous, pp. 218-238. Oxford: Oxford University Press.

Peirce, C. Sanders. 1905. Issues of pragmaticism. *The Monist* 15: 481-499.

Plantinga, Alvin. 1993a. *Warrant: The Current Debate*. Oxford: Oxford University Press.

————— 1993b. *Warrant and Proper Function*. Oxford: Oxford University Press.

Pylyshyn, Zenon. 1999. Is vision continuous with cognition? *Behavioral and Brain Sciences* 22: 341-365.

Raftopolous, Athanassios. 2019. *Cognitive penetrability and the epistemic role of perception*. London: Palgrave Macmillan.

Russell, Bertrand. 1920. *The nature of inference*. In *The collected papers of Bertrand Russell: Essays on language, mind, and matter slater, 1919-26*, eds. Bernard Frohmann and John Slater, editors, Ch. 15. Unwin Hyman: London.

Siegel, Susan. 2012. Cognitive penetrability and perceptual justification. *Nous* 46: 201-222.

————— 2013a. Can selection effects influence the rational role of experience? In *Oxford Studies in Epistemology* Vol. 4, ed. Tamar Gendler, 240-270. Oxford: Oxford University Press.

————— 2013b. The epistemic impact of the etiology on experience. *Philosophical Studies* 162: 697-722.

————— 2013c. Reply to Fumerton, Huemer, and McGrath. *Philosophical Studies* 162: 749–757.

————— 2017. *The rationality of experience*. Oxford: Oxford University Press.

————— 2019. Inference without reckoning. In *Reasoning: New essays on theoretical and practical thinking*. In eds. Brendan Balcerak Jackson and Magdalena Balcerak Jackson, 15-3. Oxford: Oxford University Press.

Silins, Nicholas. 2016. Cognitive penetration and the epistemology of perception. *Philosophy Compass* 11: 24-42.

Stokes, Dustin. 2012. Perceiving and desiring: a new look at the cognitive penetrability of experience. *Philosophical Studies* 158: 479-492.

Stroud, Barry. 1979. Inference, belief, and understanding. *Mind* 88:179-196.

Sudduth, Michael 2018. Defeaters in Epistemology. *Internet Encyclopedia of Philosophy*. http://www.iep.utm.edu/ep-defea/ Accessed 9 August 2019.

Teng, Lu. 2016. Cognitive penetration, imagining, and the downgrade thesis. *Philosophical Topics* 44: 405-426.

Tucker, Chris. 2010. Why open-minded people should endorse dogmatism. *Philosophical Perspectives* 24: 529-545.

———— 2011. Phenomenal conservatism and evidentialism in religious epistemology. In *Evidence and religious belief*, eds. Kelly James Clark and Raymond J. VanArragon, Ch. 4. Oxford: Oxford University Press.

———— 2013. Seemings and justification: an introduction. In *Seemings and justification: new essays on dogmatism and phenomenal conservatism,* ed. Chris Tucker, 1-29. New York: Oxford University Press.

————— 2014a. If dogmatists have a problem with cognitive penetration, you do too. *Dialectica* 68: 35-62.

————— 2014b. On what inferentially justifies what: The vices of reliabilism and proper functionalism. *Synthese* 191: 3311-3328.

————— Forthcoming. Dogmatism and the epistemology of covert selection. In *Reason, bias, and inquiry: new perspectives from the crossroads of epistemology and psychology*, eds. Nathan Ballantyne and David Dunning. Oxford: Oxford University Press.

Vahid, Hamid. 2014. Cognitive penetration, the downgrade principle, and extended cognition. *Philosophical Issues* 24: 439-459.

Wu, Wayne. 2013. Visual spatial constancy and modularity: does intention penetrate vision? *Philosophical Studies* 165: 647-669.

————— 2017. Shaking up the mind's ground floor: the cognitive penetration of visual attention. *Journal of Philosophy* 114: 5-32.