

# On the Phenomenology of Thought

*Joseph Levine*

## 1 Introduction

Phenomenal consciousness is usually captured by Nagel's (1974) phrase "what it's like." A phenomenally conscious experience is the kind of state there is something it is like to occupy, and the phenomenal character of the experience is precisely what it is like. One way to put it is this. Phenomenal consciousness involves two fundamental components: subjectivity and qualitative (or phenomenal) character. There being something at all it's like to have an experience manifests subjectivity. Subjects of experience are the entities *for whom* the experience is like what it's like. Qualitative, or phenomenal character is the complex of features presented to the subject of experience. That subjectivity at all should exist in the world is, to my mind, itself a major puzzle, independently of the puzzles that surround the particular features presented to us as phenomenal characters.

The standard examples of phenomenal states involve sensation, and maybe emotion too. Feeling pain, seeing red, tasting salty, and perhaps also feeling angry or sad, are thought to be examples of states with phenomenal character. Mental states that lack phenomenal character are all those states, including non-occurrent beliefs and desires, that are classified as unconscious. But what about occurrent cognitive states? I now consciously, occurrently believe that I am typing on my computer, that the sun is shining (for a change), and I'm also wondering what's for dinner tonight. I certainly seem to be conscious of these thoughts; or, to be more careful, I certainly am consciously thinking these thoughts. (Just what it is I am conscious *of* while occurrently thinking that P is yet to be decided.) But do these cognitive states possess phenomenal character? Is there something it is like to be having these thoughts?

Let us call the doctrine that these questions have an affirmative answer "CP," for "Cognitive Phenomenology." CP has been strongly defended by a number of philosophers<sup>1</sup>, and also denied by others<sup>2</sup>. My aim in this paper is twofold: first, I want

<sup>1</sup> See Goldman (1993), Horgan and Tienson (2002), Pitt (2004), and Siewert (1998).

<sup>2</sup> See Lormand (1996) and Nelkin (1989).

to distinguish various possible versions of CP, ranking some as stronger than others; second, I want to examine the standard arguments in favor of CP and see just which versions of the thesis they support. I will conclude that though the standard arguments do make plausible at least a relatively weak version of CP, they do not support stronger versions.

## 2 Two arguments for CP

There are two basic pro-CP arguments I want to consider: what I'll call "the self-knowledge argument" and "the phenomenological argument." The first appeals to phenomenal character as the only, or at least the best, explanation of how we have the particularly intimate knowledge of our conscious cognitive states that we do. The second appeals to our own experiences to demonstrate the existence of cognitive phenomenology.

Let's consider the self-knowledge argument. Of course, I have a multitude of beliefs and desires at any one moment, and clearly, in one sense I am not aware of them all at once. These are the unconscious cognitive states. But when I make one of them conscious—when I consciously entertain a thought—I am automatically, it seems, aware that I am thinking it. What's more, this immediate awareness of what I'm occurrently thinking constitutes knowledge. If I know anything, it seems, I know what I'm now consciously thinking. So what can explain this sort of self-knowledge?

The CP advocate answers that we know what we're thinking by experiencing the phenomenal character of consciously thinking it. Just what this means, and how it explains what requires explanation here—indeed, just what it is that requires explanation here—are all issues I will take up in some detail as we proceed. For now, I just want to get the arguments on the table.

The phenomenological argument is best seen as a reply to a standard anti-CP ploy. Consider again my occurrent thought that the sun is shining. If I take stock of what it's like for me now, what do I find? Of course there are all the sensory states with their phenomenal characters. But when it comes to the thought itself, one might argue, all the phenomenal character there is to be found is that which belongs to the auditory and visual imagery connected with the thought—the sound of "inner speech" and perhaps some visual images as of the sun shining, etc. Of course, no one doubted that imagery of this sort possessed phenomenal character (unless you were skeptical of phenomenal consciousness in general, a view we're not considering here), so if CP is a substantive thesis, it has to entail the presence of phenomenal character that goes beyond such imagery. But, the objection goes, there isn't any such phenomenal character to be found.

In response, the CP advocate points to phenomena like ambiguities, the difference between rehearsing a sentence in a language one doesn't understand and "hearing" it in inner speech when one does understand it, the moment of "getting" what a garden path sentence means, and the like. The general idea is this. If all phenomenology is

sensory phenomenology, then there shouldn't be a difference in what it's like to have the very same sound sequence imagined in your head when it means one thing rather than another, or nothing at all. For instance, as a young boy I was taught various blessings in Hebrew which I learned to recite at the proper times without understanding what they meant. As I got older I learned quite a bit of Hebrew, and now when I say, or imagine one of these blessings I know what it means. (Alas—I liked them better when I didn't know what they meant.) The CP advocate will ask: Isn't there clearly a difference in what it's like to hear the blessing in your head when you understand it from what it's like when you hear it without understanding? If so, then this difference is attributable to the phenomenal character of the grasping of the thought or content in question, and not merely that which comes with the sound of the words that express that content.<sup>3</sup>

So these are the two principal pro-CP arguments I want to consider. In order to evaluate the arguments, it's helpful to specify in a little more detail what the anti-CP position is. While obviously there are lots of different views one might have on the nature of occurrent thought, or any thought at all, I am going to take the following as our "null hypothesis," the default anti-CP position. Let's call it "NPFR," for "Non-Phenomenal Functional Representationalism." On this view, as on any functionalist-representationalist view, the mind is a representational system, with representational states embodied in physical configurations in the brain. Thinking is a matter of tokening certain "mentalese" sentences and processing them in various ways. The different attitudes are constituted by different functional relations to the relevant mentalese sentences, and the different contents toward which one can take an attitude are determined by the semantic properties of the mentalese sentences. These semantic properties are determined by causal or nomic relations to the world, and also (perhaps) by functional relations among the sentences themselves. What NPFR adds to standard functionalist-representationalism is this stipulation: While there is (or may be) phenomenal character experienced as a result of occupying certain perceptual states, there is no corresponding phenomenal character experienced as a result of occupying even occurrent cognitive states. What distinguishes occurrent from non-occurrent cognitive states is completely exhausted by functional features, such as whether or not a certain sentence is currently tokened in a certain location. So the question now is whether the two arguments, either singly or together, show that NPFR is wrong.

Before proceeding to look at the arguments in detail, a note about terminology is in order. Notice that I framed NPFR in terms of a thought's being "occurrent" or "non-occurrent," not "conscious" or "unconscious." The problem is that for some (see Pitt, this volume), the term "conscious" is appropriate only for states that have phenomenal character. Obviously, CP would be vacuous on this view if it applied only to "conscious" thoughts. Others, notably Block (1995), distinguish "access consciousness" from "phenomenal consciousness," and thus could frame the question as whether

<sup>3</sup> Siegel (2007) calls this form of argument "the method of phenomenal contrast."

any access-conscious thoughts are also phenomenally conscious. On Pitt's understanding of "conscious," the null hypothesis is that thoughts are themselves never conscious, only their sensory accompaniments. On Block's understanding of "conscious," the null hypothesis is that access-conscious thoughts are never phenomenally conscious. I will try to stick to the neutral characterization of our target phenomena as "occurrent thoughts" in what follows. As we will see below, however, the issue is not purely a matter of terminology, but has substantive consequences as well.

### 3 Evaluating the two arguments

The following lengthy quote from Pitt (2004) presents the self-knowledge argument and also serves to fill out CP somewhat in the process:

Normally—that is, barring confusion, inattention, impaired functioning, and the like—one is able, consciously, introspectively, and non-inferentially (henceforth, "Immediately") to do three distinct (but closely related) things: (a) to distinguish one's occurrent conscious thoughts from one's other occurrent conscious mental states; (b) to distinguish one's occurrent conscious thoughts each from the others; and (c) to identify each of one's occurrent conscious thoughts as the thought it is (i.e. as having the *content* it does). But (the argument continues), one would not be able to do these things unless each (type of) occurrent conscious thought had a phenomenology that is (1) different from that of any other type of conscious mental state (proprietary), (2) different from that of any other type of conscious thought (distinct), and (3) constitutive of its (representational) content (individuating). That is, it is only because conscious thoughts have a kind of phenomenology that is different from that of any other kind of conscious mental state that one can Immediately discriminate them from other kinds of conscious mental states; it is only because type-distinct conscious thoughts have type-distinct phenomenologies (of the cognitive sort) that one can Immediately distinguish them from each other; and it is only because a conscious thought that *p* has a phenomenology that constitutes its (representational) content that one can Immediately identify it as the thought it is. Hence (the argument concludes), each type of conscious thought has a proprietary, unique phenomenology, which constitutes its representational content. (pages 7–8, emphasis in original)

According to Pitt, CP (his version of it) involves three elements: that cognitive phenomenology is *proprietary*, *distinctive*, and *individuating*. That it's proprietary means that it involves its own kind of phenomenology; it's not sensory in any way. That it's distinctive means that there is a different one for each thought. These two features are fairly clearly explained in the quote. But what it means to say that the phenomenal character of an occurrent thought "constitutes its representational content" is not so clear, and I will return to it below. For now, however, I just want to evaluate the general argument that we must appeal to CP to explain what Pitt calls our "Immediate" knowledge of what we're occurrently thinking.

Pitt goes through a number of possible alternative explanations, finding them all wanting. The alternative I want to consider specifically is the one that naturally follows from NPFR, our null hypothesis. What it is to have knowledge of what one is thinking

is to token a mental representation—a mentalese sentence—that expresses the fact that one is thinking what one is thinking. What makes this Immediate knowledge, in Pitt's sense, is the fact that this sentence tokening is not the result of an inferential process, but rather an immediate causal result of the first-order thought state itself (together with some functionally characterizable internal monitoring process). It's because of the reliability of the relevant process yielding the higher-order sentence expressing the fact that one is thinking a certain content that it counts as knowledge. If this explanation is adequate, then we don't need to appeal to the thought's phenomenal character to explain how we know—Immediately—that we're thinking it.<sup>4</sup>

Pitt considers this alternative, but presents two objections to it. I'm not awfully sure that I understand his objections, so I'm going to quote him and then provide my interpretation. If I'm right, it will turn out that the objections don't succeed. The first objection is captured in the following passage:

But to think that *t* is the thought that *p* while *t* is occurring—even because *t* is occurring—is not to *identify it* as the thought that *p* in the sense at issue in this paper. Introspective identification of occurrent conscious thoughts is analogous to perceptual identification of objects and introspective identification of sensations: it is a form of knowledge *by acquaintance*. Such identifications have the canonical form *this [that] is (an/the) F*; they require simple acquaintance, in the relevant mode, with the object identified. That is, the object identified—“this”—must be *experientially discriminated by the perceiver from its environment*. And, as pointed out above, this requires that the object appear to one in some determinate way, and that one be attending to it. One cannot, say, visually identify a thing as a dog (see that it is a dog) unless one has an attentive discriminative visual experience of it—that is, unless one *simply sees* it (in the sense of section 2). Merely to think *this animal is a dog* when a dog is within visual range and is causing one to have a visual experience of it and to think that it is a dog, is not to *see that* it is (visually identify it as) a dog. Likewise, merely to think that *s* is a pain when *s* is a pain, is occurring, and is causing one to think that it is a pain, is not to *feel that* it is (introspectively identify it as) a pain. And merely to think that *t* is the thought that *p* when *t* is the thought that *p*, is occurring, and causing one to think that it is the thought that *p*, is not to *grasp that t* is (introspectively identify it as) the thought that *p*. (page 19, emphases in original)

I find this passage quite interesting for what it suggests concerning how to think about just what phenomenal character is, and I will return to this theme below, when I distinguish the various versions of CP. But for now, I just want to see what force this objection has against the advocate of NPFR. As far as I can tell, not much. The reason is that so much is packed into the notion of Immediate self-knowledge that it appears to beg the question. This is the point, mentioned above, where what seemed merely a terminological issue becomes more substantive.

<sup>4</sup> See Nichols and Stich (2003) for an account along these lines. Nelkin (1989) also endorses a view like this, though he explicitly refrains from committing himself to the claim that meta-beliefs formed this way constitute knowledge.

Pitt characterizes “Immediate” self-knowledge as “conscious.” If by “conscious” one only means “occurrent,” or “access conscious,” then there’s no harm in it. But if one insists that our “conscious” knowledge of what we’re currently thinking is a matter of perceptual-like acquaintance—comparable to how I know what I’m seeing or feeling—then I guess it must involve phenomenal character. This is tantamount, however, to just describing it as phenomenally conscious. But then we don’t need to appeal to phenomenal character to explain this kind of knowledge, since the alleged explanandum is already described as a phenomenal state. If, on the other hand, we refrain from building phenomenal character into the description of the kind of self-knowledge we’re trying to explain, then Pitt’s objection seems to disappear. After all, it certainly seems as if one can hold a position to the effect that we have immediate (perhaps not “Immediate” if that entails perceptual-like appearance) knowledge of what we’re thinking—it’s just, as it were, magically available to us, in that we don’t discern any process of discovery or inference—without there being anything like a phenomenal appearance, as we have with sensory experience. To demand an explanation of this richer, appearance-like knowledge is, as I said, to beg the question.

The second objection is captured in the following passage ( $t'$  is the higher-order thought to the effect that first-order thought  $t$  has the content  $p$ ):

...  $t'$ 's consciousness is supposed to make the content of  $t$  Immediately knowable. But this is because the content of  $t'$  is that the content of  $t$  is  $p$ . Hence, if the consciousness of  $t'$  is not sufficient to make its content Immediately knowable—i.e. to make it Immediately knowable that it says that the content of  $t$  is  $p$ —then it is not sufficient to ground Immediate knowledge of  $t'$ 's content. Hence, conscious occurrence of  $t'$  must, if it is to be sufficient to ground Immediate knowledge of the content of  $t$ , be sufficient to ground conscious knowledge of its own content as well. Since the theory under consideration denies this, it is false. (page 20)

I don’t think this objection works either. To say why, I need to introduce a distinction between *implicit* self-knowledge and *explicit* self-knowledge. Explicit self-knowledge is what we have when we explicitly formulate a metacognitive thought, such as “I believe that San Francisco is a beautiful city.” Someone asks me, say, if I believe San Francisco is beautiful, and I reply that yes, I do. This seems to be one of those special, first-person ways of knowing what I believe. I don’t have to infer it from observing my behavior, and it certainly doesn’t seem as if I wait to see what inner speech goes through my head and then interpret that and infer what I believe. I just seem to know.

What I’m calling “implicit” self-knowledge, however, is not the result of any explicit formulation or reflection. Rather, it’s the knowledge that seems to come with the very thinking of the thought itself. (Perhaps it’s not really apt to call this “knowledge” of what I believe; that might be only appropriate to say of the explicit self-knowledge case. But I’m going to continue to use the term “implicit self-knowledge” for ease of exposition.) As I think, and let’s imagine this involves some inner speech, I understand what it is I’m “saying to myself,” and in that sense know what it is I’m thinking. No explicit metacognitive proposition has been formulated. I think to myself,

as I gaze at the view from Dolores Park, “This sure is a beautiful city.” As I think it, it seems, I know what it is I’m thinking. This kind of knowledge is implicit self-knowledge.

Now the objection in the last passage quoted seems to assimilate these two forms of self-knowledge. The problem is supposed to be that unless I Immediately know the content of  $t'$  I can't Immediately know the content of  $t$  itself. But, on the NPFR account, Immediately knowing the content of  $t'$  would require yet another thought,  $t''$ , and then we're off and running. But that's not the right picture. On NPFR one gets implicit self-knowledge for free. All that's required is that one thinks in one's language of thought, mentales. To implicitly know what one is thinking is just to think it with understanding. But it's a mistake to view thinking with understanding as a matter of interpreting one's own thoughts. No, one just thinks them. If they have the requisite semantic properties, then that's all there is to it.

Explicit self-knowledge, on the other hand, does involve formulating another thought—a thought to the effect that one is having the first-order thought. But for that to work, all we need is for the subject to token a higher-order representation that means that the lower-order representation has content  $p$ . Here, the subject is explicitly reflecting on what she believes (thinks, desires, etc.) and does assign an interpretation. But doing this—thinking, interpreting, etc.—is, on NPFR, just a matter of tokening the relevant mentales sentences with the right semantic properties. We can put it this way. To know (explicitly) is just to occupy another cognitive state, and that, on this theory, is just to token the right representation in the appropriate circumstances. To explicitly know thought  $t$ 's content is to think another thought,  $t'$ , whose content is that the content of  $t$  is  $p$  and is itself implicitly known. On NPFR, it's tokening and processing sentences all the way down.<sup>5</sup>

Let's turn now to the phenomenological argument. We are asked to reflect on various possible experiences and see if we find that there is a difference in what they are like. So, imagine that you are having a certain course of visual and auditory imagery. Keeping that sensory/imagistic phenomenology constant, imagine two scenarios: one in which you are saying a sentence you understand to yourself, and another in which you don't understand it. If you think there's a difference in what the two experiences are like for you, then that difference—given we've held the sensory/imagistic phenomenology constant—must be attributable to the phenomenological contribution of grasping (or not) the relevant content.

<sup>5</sup> In Pitt (this volume) and personal communication, Pitt objects that the sort of implicit self-knowledge I attribute to the subject—i.e. just understanding what she's thinking by tokening, in functionally appropriate circumstances, the relevant sentence of mentales—doesn't capture what it is to consciously know what one is thinking. Again, I agree, so long as one understands “consciously” in as rich a fashion as Pitt does. But then the anti-CP advocate will deny that we have such conscious self-knowledge in the first place, and so therefore there's nothing of the kind to explain. I will return to the question of just how to characterize the explanandum below.

Of course, one possible line of reply is that in such cases the sensory/imagistic phenomenology is not actually held constant. To see what I mean, consider the debate over perceptual modularity. Fodor (1983) and Pylyshyn (2001) defend a strong modularity thesis on which there is no cognitive penetration from higher cognitive processes into the deliverances of the visual module. An alleged counterexample to this thesis is the Gestalt switch phenomenon, the “duck–rabbit” being one of the best well-known examples. The point is that deciding to see it as a rabbit—a higher-level conceptual process—clearly affects how it appears, what it looks like. Hence, there seems to be genuine cognitive penetration from above.

There are a number of ways to defend the modularity thesis against this sort of example, but one way in particular, mentioned by Pylyshyn (2001), is relevant to our case. Pylyshyn discusses research that shows that when you decide to see the figure as a rabbit, say, rather than a duck, your eyes actually focus on different points in the figure from where they focus when you intend to see it as a duck. In a sense, the higher-level process is influencing the way it looks, but not by direct cognitive penetration. Rather, one’s decision affects the precise character of the stimulus presented to the eyes, which then, in a strictly bottom-up fashion, produces a difference in how the figure looks.

Whether this move works for all examples of Gestalt switches I don’t know. But it seems to me a similar move could be made by the advocate of NPFR in response to the phenomenological argument. That is, one might argue as follows. When one understands a sentence performed in inner speech—or understands it as read one way rather than another—this affects the prosody, or some other perceptually manifested feature of the imagined sound stream. So the difference in what it’s like between understood inner speech and non-understood inner speech might just be a sensory difference after all. When you understand what you’re saying to yourself—or understand it this way rather than that way—you actually say it differently, and this is the phenomenological difference you detect. It’s not a direct phenomenological reflection of grasping the content, of the understanding itself.

I don’t know how far this move can go in accounting for the phenomenological data. I’m inclined to think it’s a promising avenue to explore, but in the end I doubt it can do all the work needed. For now, I propose to concede to the advocate of CP that the phenomenological argument does establish that at least some version of CP holds. But which version? How should we understand CP, and what do the two arguments—the self-knowledge argument and the phenomenological argument—tell us about the nature of cognitive phenomenal character? This is the topic of the next two sections.

#### 4 Distinguishing stronger and weaker versions of CP

At the beginning of the paper I characterized phenomenal character as what it is like to occupy a mental state when there is something it is like to occupy it. So



the phenomenal character of a pain state is the precise way that pain feels, and the phenomenal character of a visual sensation is the precise way things in one's visual field look. While this combination of a very general, and rather abstract characterization of phenomenal character (the "what" in "what it's like") with specific sensory examples suffices to provide substance to the notion when dealing with sensory experiences, it would be nice to have something more to say if we are to attribute phenomenal character to cognitive states. True, we can say that there is something it is like to occurrently believe that *p*, and the phenomenal character is just what it's like. But given that sensory phenomenology constitutes the only really uncontroversial example of phenomenal character, it would be helpful to find some more concrete characterization of the phenomenon that might plausibly capture both sensory and cognitive phenomenology.

I don't have anything precise to offer in this regard, but it seems to me that the following helps to flesh out the notion somewhat. In fact, what I have in mind is closely related to David Pitt's remarks, quoted above, concerning the relation between phenomenal character and acquaintance. To be phenomenally conscious, I want to say, is to be "appeared to." It is for the conscious subject to be experientially presented with a determinate object (or objects) with determinate qualities. I think this is the notion that Kant was after when he spoke of "intuition." Intuitions for Kant, as I understand it, are what provide the understanding with the concrete, singular material to which to apply concepts. Phenomenal character, then, is not just "what it is like," but more fundamentally, it's "what is presented" in experience.<sup>6</sup>

With this, admittedly still vague, characterization of the phenomenal in mind, there are two, orthogonal distinctions I want to introduce concerning the notion of cognitive phenomenology; both distinguish between a weaker and a stronger version of CP, though along different dimensions. The first distinction can be described using the Kantian terminology just introduced. On the one hand, one might hold that all phenomenal experiences, all intuitions, involve what Kant called a "sensory manifold," the forms of which are Space and Time. On this view, while all phenomenal character is essentially sensory—involving the distribution of sensible features (color, shape, sound, etc.) in space and time (or sometimes only in time)—cognitive states make a distinctive contribution to the precise phenomenal character experienced. The idea is that what one is currently thinking/believing/desiring affects the way a particular sensory manifold appears, so for two experiences with identical purely sensible features but different thoughts, the subjects will experience the two sensory manifolds in different ways. However, all phenomenal character is grounded in sensory

<sup>6</sup> This characterization of phenomenal character assumes a kind of representationalism about phenomenal character, though not of a reductionist variety. The alternative is to view phenomenal character as a kind of "mental paint," an intrinsic quality of experiential states that has no inherent intentional or representational features. Since both I and the CP advocates discussed in this paper reject the mental paint view, it's safe to set it aside without argument for present purposes.

phenomena, so that cognitive states can only be consciously experienced through their effects on what sensorily appears to the subject.

One might put it this way. On this version of CP, which I'll call "impure CP," cognitive phenomenology is a matter of one's sensory experience being cognitively inflected. What one is thinking changes the way what one is perceiving (imagining, etc.) appears to one, but in the end all phenomenology involves the appearance of some sensorily presented object and its qualities. So, on this view, when you visually experience something you recognize it has a different "look" from the one it has when you don't recognize it, as does speech you understand have a different "sound" from the one it does when you don't understand it. But it's still the case that the only way to phenomenally experience a cognitive content is through its effect on some sensory presentation.

On the other hand, one might claim that there is a phenomenology of pure thought, a doctrine I'll call "pure CP." That is, independently of any sensory phenomenology, and not strictly through its effect on one's sensory phenomenology, there is just something it is like to think a thought with a certain content. Now, given my characterization above of phenomenal character as involving the presentation of some object and its qualities, how would this apply in the case of purely cognitive phenomenology? Here's one way to put it, and, again, I think this fits fairly well with Pitt's characterization of cognitive phenomenology when he describes it as a kind of acquaintance: the content of a thought is presented to the subject thinking it—it appears to the subject—but in a non-sensory way. Perhaps this is what Kant had in mind when he speculated about the possibility of "intellectual intuition," the concrete, singular presentation of a purely cognitive, conceptual object.

Between pure CP and impure CP, which position is Pitt defending? I'm inclined to think it's pure CP, since the notion of a "distinctive" cognitive phenomenology seems to fit that position better.<sup>7</sup> However, since on the impure version the admixture of cognitive content creates distinctions among sensory phenomenal characters that wouldn't exist otherwise, one can imagine calling that contribution a kind of "distinctive" phenomenology as well, even though it can't, as it were, stand on its own. But whichever version Pitt intended, what interests me here is which version is supported by the two arguments, the self-knowledge and phenomenological arguments. Before turning to that question, though, let me introduce the second distinction.

The question on which this distinction turns is whether cognitive phenomenology shares with sensory phenomenology a feature I'll call "transparent content." What I mean is this. When I have a conscious visual experience of, say, a ripe tomato sitting on a kitchen counter, a proper characterization of what it's like for me to have this experience makes essential reference to how the space around me appears. Round shapes and red colors constitute the look, and there doesn't seem to be any space, as it

<sup>7</sup> I think this also fits Siewert's defense of the claim that what he calls "noniconic" thought has phenomenal character.

were, between what it's like for me and how the world is being presented to me. This characterization of visual phenomenology of course coheres with my account of phenomenal character at the beginning of the section.

If cognitive phenomenology has transparent content (call this CPTC), that means that what the cognitive state is about, what it's representing, constitutes the "look" as it were of the cognitive state. What appears to me when phenomenally experiencing a belief, say, is the belief's content, how I'm believing the world to be, in just the way that what appears to me when I see a ripe tomato is how I'm visually experiencing the world to be. If cognitive phenomenology has transparent content, then it can serve the epistemological function Pitt assigns it; it can acquaint me with what I'm thinking.

Suppose, however, that cognitive phenomenology lacks transparent content—let's call this having "opaque content" (CPOC).<sup>8</sup> What would that mean? I can imagine two alternatives. The first is the "mental paint" view of cognitive phenomenology.<sup>9</sup> On this position, it turns out there is just something it is like to believe that *p*, but what it's like is an intrinsic feature of the believing experience and has no essential connection—though presumably a causal one—to the content believed. While I don't see any outright inconsistency in holding a mental paint view of cognitive phenomenology and rejecting it for sensory phenomenology, it's hard to see what motivation there would be for doing so. In particular, the idea that phenomenal character is a matter of how something determinate appears to a subject is inconsistent with the mental paint account, so adopting the mental paint view for cognitive phenomenology would mean abandoning the general account of phenomenal character presented above.

The second alternative preserves the idea that phenomenal character is "presentational" in nature. However, rather than presenting a content, what is presented is one's own mental state; on a representational theory of mind, what is presented is the underlying mental representation that is the immediate relatum of the cognitive attitude in question.<sup>10</sup> The picture I have in mind is this. Consider again the anti-CP position mentioned earlier, on which there is no cognitive phenomenology at all, but instead only conscious awareness of sensory experiences associated with inner speech, imagery, and the like. The opaque content version of cognitive phenomenology I'm imagining here allows that there is more than the sensory experience of inner speech and imagery, but claims that this something more is still analogous to the sensory phenomenology of inner speech in that it is a kind of cognitive "hearing" of underlying mental representations. Just as I am presented with auditory imagery of spoken

<sup>8</sup> I realize that by using this terminology I am inviting confusion between this distinction and the one between extensional and intensional contexts. However, the metaphorical character is significantly similar in both cases. With extensional contexts, where substitution of co-referentials is allowed, the idea is that we somehow see directly through the representation to the object itself. Opaque contexts are those in which we can't see through the representation. I mean something similar here, except that the metaphor is being taken a little more seriously in my use of the transparent/opaque distinction than it is in the standard use concerning extensional and intensional contexts.

<sup>9</sup> See note 6 above.

<sup>10</sup> Hence, the appropriateness of the term "opaque," as noted in footnote 8 above.

speech when thinking out loud, I can also be presented with cognitive imagery of underlying speech when consciously thinking. Whether it's genuine inner speech, or non-sensory experience of a conscious occurrent state, all conscious thought, on this view, is a matter of being presented with what's going on in one's own mind.

We now have two apparently independent dimensions along which to distinguish versions of CP: pure vs. impure, and CPTC vs. CPOC. This gives us four possible versions of CP altogether. Now as I see it, pure CP is a stronger hypothesis than impure CP, since the latter restricts the realm of the phenomenal to the sensory, though allowing fine distinctions in phenomenal character that have a cognitive source. Cognitive phenomenology, on impure CP, must always find a sensory manifold on which to appear to the subject, whereas it is often experienced independently of sensory phenomenology on pure CP.

Similarly, CPTC—the idea that fairly abstract, non-sensible contents can be objects of direct acquaintance—seems a stronger position than CPOC, the view that all cognitive phenomenology adds is a conscious acquaintance with our own mental representations. One can see this from the characterization of CPOC above, and its relation to NPFR, the null hypothesis that there is no cognitive phenomenology at all. On the latter view, phenomenology is about how our senses present the world to us. It is restricted to those features that the senses can depict.<sup>11</sup> So while there is a sense in which we are directly, consciously acquainted with colors and shapes, say, we are not so acquainted with non-sensible properties. It's not that our thoughts don't represent mind-independent reality on this view; it's not a matter of skepticism. Sure, we think, and in that sense we cognitively grasp mind-independent contents. But the point of denying cognitive phenomenology is to restrict our conscious access to the sensory manifestations of our mental representations; to leave us consciously behind the veil of inner speech and imagery.

CPOC maintains the veil that separates us from our cognitive contents, it just allows that we have a kind of phenomenological access to underlying representations that are not sensorily manifested.<sup>12</sup> The veil in question is not constituted only by our sensory

<sup>11</sup> Of course, which worldly features the senses can depict is controversial. A standard empiricist view would restrict the senses to depicting clearly sensible properties such as colors, shapes, textures, motion, etc. Siegel (2006) argues for a more expanded range of properties. How her view concerning the contents of sensory experience relates to the debate over cognitive phenomenology is an interesting question, but one I can't address in this paper.

<sup>12</sup> Now of course there is considerable controversy over just what a content is. I would like to avoid dealing with that issue if at all possible, just acknowledging that one basic division between theories of content is between Russellian and Fregean views. For Russellians, contents are something like states-of-affairs, actual components of the world (or constructed out of such components). On the Fregean view there is an intermediary between the mind, and representations generally, and the states-of-affairs that constitute truth conditions. Contents are senses, or modes of presentation of states-of-affairs. The idea that we have "Immediate" access to contents, that contents appear to the conscious subject, seems to make more sense on a Fregean view of content than on a Russellian one, but the issues here are complex. Anyway, as I said, I can't go into the extremely difficult question of what a content is in this paper, so I'll leave the matter with just one more brief comment on Fregean contents.

presentations, but also by non-sensory ones; the point is, however, a veil exists on both positions. CPTC, on the other hand, claims that we have as much acquaintance with the actual content of our thought as we do with the content of our sensory experiences. Clearly, then, CPTC is a greater deviation from NPFR than is CPOC.

What I want to do in the next section is return to the two arguments, the self-knowledge and phenomenological arguments, and determine which of the versions of CP—pure or impure, TC or OC—they support. Just what do these arguments buy us in the way of cognitive phenomenology? I will argue for two claims. First, the phenomenological argument only buys you the weaker alternative of both distinctions, namely, impure CPOC. Second, understood in the way we have so far, the self-knowledge argument buys you very little at all, since, as I argued above, NPFR, without appeal to any cognitive phenomenology, can explain the relevant self-knowledge. However, understood another way, it's possible it gets you at least impure CPTC. Whether or not it does depends on figuring out just what it is we know about what we're thinking "from the inside," and I conclude with some very inconclusive, speculative remarks on that topic.

## 5 Which version is supported by the two arguments

So let's turn to the phenomenological argument. Remember, it enters the dialectic as a response to someone who claims not to notice anything but sensory phenomenology in her stream of consciousness. The argument addresses this objection by noting many circumstances—whether it be ambiguous sentences or visual figures, the difference between rehearsing a sound sequence in one's head while understanding and not understanding, or many other such examples—in which the cognitive state one occupies makes a difference to what hearing, seeing, or otherwise sensing something is like. I mentioned above that there are moves to be made by the advocate of NPFR in reply, but I myself think there is something to the phenomenological argument. The question is, what does it buy you in the way of cognitive phenomenology?

When it comes to the distinction between pure and impure CP, it seems obvious that this argument can't support more than impure CP. After all, all the examples

The comment is that on at least one conception of mode of presentation (MOP) this difference between CPTC and CPOC that I'm emphasizing now disappears. The idea is this. If we think of a MOP as what explains how we can be ignorant of identities, how we can rationally believe that Hesperus appears in the evening but Phosphorus doesn't, then on a representational theory of propositional attitudes we can identify MOPs with mentalese representations. Venus, the object, is presented to me one way by thinking of it with my mentalese equivalent of "Hesperus," and another way by thinking of it with my mentalese equivalent of "Phosphorus." The point is that, for the representationalist, there is no reason to posit some entity, a sense, that is distinct from the mentalese word and is presented to the mind when thinking of the object using that word. The MOP can just be the word itself. It "presents" the object by referring to it. End of story. If we take this view of Fregean content (which, admittedly, is a very deflationary account of content), then both CPTC and CPOC allow that we have Immediate access to content. (I thank Louise Antony for making this point in discussion.)

involve ways of experiencing visual appearances, sound streams, and the like. The argument is addressing someone who doesn't find anything but sensory phenomenology in her conscious life, and it pushes her to notice distinctions among these sensory appearances that can only be accounted for by appeal to cognitive penetration.<sup>13</sup> Fair enough. Still, this doesn't get us anything like pure CP. It's still possible, for all this argument demonstrates, that the only way for a cognitive content to make itself appear to a conscious subject is through affecting the way some sensory manifold appears. That's impure CP.

How about CPTC versus CPOC? Here too, I think the phenomenological argument only buys you the weaker version, CPOC. What we can tell from the difference between what it's like to see Hebrew script, say, with understanding and what it's like to see it without understanding, is just that: that there is a difference in what it's like. For all that's manifest to us, merely by noting this difference, what we are responding to is the difference between the representational states we're occupying in the two circumstances. I'm not making the move here that's sometimes made by saying that the only phenomenological difference noticed in such cases is a general feeling of understanding, as opposed to a feeling of puzzlement. On that view, which an advocate of NPFR might accept, there is a kind of emotional phenomenal character added to the strictly sensory one, but nothing that is either distinctive or individuating of the cognitive state involved in understanding the script. No, I'm allowing that the cognitive state contributes in a way that is distinctive of cognition and specific to this cognitive state. But what makes this cognitive state the state it is may very well be the identity of the representation that constitutes it. Yes, we understand, and grasp what the script means, but what it is, on this view, to understand what something means is to token a representation that means what the script means. We now add that there is something distinctive that it is like to occupy such a state.<sup>14</sup>

So far I've argued that the phenomenological argument itself only supports impure CPOC, the weakest of the versions of CP that we've seen. Let's turn now again to

<sup>13</sup> Using this term of course brings to mind Fodor's (1983) modularity thesis. In fact, these sorts of examples must be explained away on his view in one way or another. One possibility, discussed by Fodor, is that they represent post-modular processing. The other possibility, also discussed by Fodor, is that certain conceptual representations are available within the perceptual modules. Notice that if we go with this line, then there ought to be a clear distinction between those thoughts capable of penetrating sensory appearances and those that can't. I can't go into this here, but it seems to me that there are such limits, and therefore the phenomenological argument can't show that there is cognitive phenomenology for all occurrent thoughts.

<sup>14</sup> Charles Siewert, present when an earlier version of this paper was delivered at Cal State LA, objected that what he is aware of when understanding a sentence is its content—what it means—not any mentalese sentence. I'm not sure I totally understand his objection, but if I do understand him correctly, I think there are two replies to make. First, "aware" in this context is opaque. What I mean is that what I am aware of is not presented *as* a mentalese sentence, but as what the public language sentence means. However, what might be the case for all that is that one's state of understanding *just is* tokening the right mentalese sentence in the right way. The awareness here, the phenomenal character, just is the conscious reflection of this representational state. We never get closer to content than that.

Second, the consideration Siewert brings up might be a reflection of one of the two ways of taking the self-knowledge argument mentioned above. In that case, I'll address it below.

the self-knowledge argument. If the question is whether we can account for how we have non-inferential, immediate, and reliable knowledge of what we're thinking, then I think the discussion in section 2 above demonstrates that no version of CP is necessitated by the phenomenon of self-knowledge of content. Our null hypothesis, NPFR, does fine. However, I do think there is another way to think about self-knowledge, a conception on which the connection to CP, and indeed to CPTC, is much stronger.<sup>15</sup>

Here's one way of getting at what I have in mind.<sup>16</sup> Consider Cartesian skeptical scenarios and their relation to knowledge claims. We all think we know all sorts of things about which Cartesian skeptical doubts can be raised. I know I'm typing on this computer now. But, the skeptic interjects, really, "for all I know" I'm really a brain in a vat. Does this entail that I don't know I'm typing on the computer? Well, I'm no epistemologist, but I will go out on a limb and say I still know it. Yes, it's subject to skeptical doubt. There is a clear sense in which everything could be as it is with me, epistemically speaking, and yet I'm a brain in a vat. I can't definitively rule it out. Yet, for all that, most of us would agree that I do know I'm currently typing on my computer.

Let's call such knowledge "dubitable" knowledge. I can doubt it, it makes sense to doubt it, I know what's being proposed when the doubt is presented, but I still know it nevertheless. Is there "indubitable" knowledge? Well, what's happening with me phenomenally—or, at least, some aspects of it—seems like a pretty good candidate. I'm now feeling a pain, or seeing red. It just seems an essential constituent of my epistemic position that I am having these experiences. I find it hard to even understand the skeptical doubt that I'm supposed to entertain here, if there is one. Yes, I might be a brain in a vat and none of what I see or hear or taste is real, but that doesn't touch my having certain auditory, visual, and gustatory experiences. Skepticism just doesn't get a grip here.

If we grant that conscious experience gives rise to at least some indubitable knowledge, then certain consequences follow. I have myself appealed to indubitable self-knowledge of phenomenal character to argue against externalist representational theories of phenomenal character.<sup>17</sup> Briefly, the argument goes like this. If phenomenal character were determined by the external, referential content of visual representations, then we could encounter Frege-cases, where two visual representations, say, pick out the same color, even though they look different. So long as phenomenal content

<sup>15</sup> It's very possible that the conception of self-knowledge I'm about to describe is what Pitt had in mind when speaking of an acquaintance-type relation. I still maintain that in the context, his appeal to acquaintance was question-begging; indeed, even at the end of the day it may be, as we'll see below.

<sup>16</sup> I owe a debt to Terry Horgan for valuable discussion of this conception of self-knowledge and its role in the cognitive phenomenology debate. See also my discussion in Levine (2001) and Horgan, Tienson, and Graham (2006).

<sup>17</sup> See Levine (2003a and 2003b). But also see Levine (2008), where I endorse a version of representationalism.

isn't identified with, or supervenient on, external content, there's no problem. But if the external content determines the phenomenal character, and the two visual representations pick out the same color, then how can they look different? That there is a phenomenal difference seems like a good example of indubitable knowledge, and my argument is that externalist representationalism cannot account for this.

Suppose we turn now to knowledge of content. Is this indubitable knowledge? Consider the Swampman hypothesis.<sup>18</sup> On some theories of content—especially teleo-semantic views, but maybe causal-covariational accounts as well—SwampJoe, my molecule-for-molecule duplicate that just appeared as a brain in a vat in deep space (just to make the case as strong as possible), doesn't genuinely think, or represent with his cognitive states, as his mental representations do not possess semantic properties. Now, I want to say this. That I'm not SwampJoe is something I know, but not indubitably. There is a clear sense in which I can coherently entertain the hypothesis that I am SwampJoe. But, initially at least, I want to say that I really can't entertain the hypothesis that I'm not thinking—I mean, really thinking—in just the way that I can't coherently entertain the hypothesis that I'm not phenomenally conscious right now. But if that's right—if indeed we have indubitable knowledge that we're thinking and what we're thinking—doesn't that cry out for explanation? It doesn't seem as if NPFR can provide that explanation. After all, the only clear example we have (if one goes this far with us “qualia freaks”) of indubitable knowledge is what is presented to us phenomenally. What's more, both NPFR and CPOC, remember, put us behind the veil of representation, with no direct acquaintance with contents.<sup>19</sup> But then the argument that thought content must be presented phenomenally, and that its phenomenal presentation is the basis for our cognitive access to it, becomes compelling. Hence, we have an argument for CPTC.<sup>20</sup>

So whether we can get CPTC out of the self-knowledge argument—indeed, given my criticisms of the argument above, whether we get any version of CP out of it—depends on what kind of knowledge we're talking about. If we do indeed have indubitable knowledge—if I can tell without possible doubt from the inside that I'm genuinely thinking, with genuine content—then I would be inclined to accept that this somehow must be a matter of phenomenal access. The problem is, unlike with the case of sensory knowledge, I don't really know how to evaluate the claim to indubitable self-knowledge of content. Suppose there were a creature that only had a non-conscious computational mechanism superimposed on a phenomenally conscious sensorium. Would it be any different for that creature than my conscious life is for

<sup>18</sup> I discuss the implications of Swampman in Levine (1996).

<sup>19</sup> But see note 12 above.

<sup>20</sup> But notice this doesn't give us an argument for pure CPTC over impure CPTC. It could still be that contents can only be phenomenally revealed to us through sensory vehicles.



me?<sup>21</sup> Can I really tell, in that particular skepticism-repelling, indubitable way, that I'm *not* that creature? I have to admit that I just don't know.

For what it's worth, here's how things seem to me phenomenologically. I don't feel that I have any special access to *what* I'm thinking about, other than by way of the representations I use to express it. Of course, the advocate of CP will press the phenomenological argument to show that I can make finer distinctions phenomenologically than those provided by the range of sensory representational vehicles I can experience. But the phenomenal distinctions argument doesn't apply to the language of thought, where all ambiguities that are distinguishable in thought are captured by distinct representations. So I don't really see how I have any more direct, or firmer grasp of what I'm thinking than that it's whatever I'm currently representing in thought.

However, *that* I'm thinking, that my thoughts are really about something, and not meaningless formal objects, does seem apparent to me in that special way. I am, to borrow a phrase that was popular among presenters at a conference I recently attended, "just emoting" here, not arguing. I don't really know if one can have indubitable knowledge that one is thinking without indubitably knowing what one is thinking.<sup>22</sup> But if you can, then I think you still don't get CPTC. For which content one is entertaining is determined by the semantic content of one's mental representations, and all one has to think with are one's representations. We're still behind the veil, on this view, but it's just that it's phenomenally apparent to us that something is on the other side.<sup>23</sup>

Emoting aside, I want to conclude as follows. Dubitable self-knowledge of content is consistent with NPFR, and so with no cognitive phenomenology at all. Indubitable self-knowledge very well might support CPTC, though whether we have such self-knowledge of content is itself dubitable. The phenomenological argument, on the other hand, does support at least one version of CP, namely impure CPOC. Though this is the weakest of the versions surveyed here, it is still a version of CP—by my lights, anyway—so in the end I do come down, though not heavily, on CP's side.<sup>24</sup>

<sup>21</sup> To forestall a possible misunderstanding here, let me say this. Horgan and Tienson (2002) argue that consideration of phenomenal duplicates supports CP. They claim that it's hard to imagine two creatures with identical phenomenal streams who nevertheless are consciously thinking distinct thoughts. I take this to be a version of the phenomenological argument, which I have already accepted. However, what's at issue now is whether this consideration also supports CPTC. I don't see that their argument does that.

<sup>22</sup> Interestingly, Dretske (2003) seems to argue for the reverse, though he wouldn't endorse my notion of indubitable knowledge. Still, he thinks we have a kind of epistemic guarantee of what we're thinking, so long as we are in fact thinking anything at all, though no such guarantee that we satisfy that condition. There are passages in Siewert's (1998) discussion that seem to rely on a distinction between knowledge that thought is going on and knowledge of precisely what is being thought that might support this idea.

<sup>23</sup> I suspect that this position as stated cannot be maintained, but that something like it can so long as we come up with the right theory of content. Material for another project.

<sup>24</sup> I want to thank Louise Antony, Terry Horgan, Michelle Montague, David Pitt, Georges Rey, Charles Siewert, Kelly Trogdon, and the audience at Cal State LA, where an earlier version of this paper was presented, for comments on an earlier draft and very helpful discussion.

## References

- Block, N.J. (1995). "On a Confusion about the Function of Consciousness," *Behavioral and Brain Sciences*, vol. 18, 227–47.
- Dretske, F. (2003). "Externalism and Self-Knowledge," in S. Nuccetelli (ed.), *New Essays on Semantic Externalism and Self-Knowledge*. Cambridge, MA: MIT Press.
- Fodor, J.A. (1983). *The Modularity of Mind*. Cambridge, MA: Bradford Books/MIT Press.
- Goldman, A. (1993). "The Psychology of Folk Psychology," *Behavioral and Brain Sciences*, vol. 16, 15–28.
- Horgan, T. and Tienson, J. (2002). "The Intentionality of Phenomenology and the Phenomenology of Intentionality," in D. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press.
- , —, and Graham, G. (2006). "Internal-World Skepticism and the Self-Presentational Nature of Phenomenal Consciousness," in U. Kriegel and K. Williford (eds), *Self-Representational Approaches to Consciousness*. Cambridge, MA: Bradford/MIT Press.
- Levine, J. (1996). "SwampJoe: Mind or Simulation?" *Mind and Language*, vol. 11, no. 1, March, 1996, 86–91.
- (2001). *Purple Haze: The Puzzle of Consciousness*. New York: Oxford University Press.
- (2003a). "Experience and Representation," in Q. Smith and A. Jokic (eds), *Consciousness: New Essays*. Oxford and New York: Oxford University Press.
- (2003b). "Knowing What It's Like," in B. Gertler (ed.), *Privileged Access: Philosophical Accounts of Self-Knowledge*. Burlington, VT: Ashgate Publishing.
- (2008). "Secondary Qualities: Where Consciousness and Intentionality Meet," *Monist*, vol. 91, no. 2.
- Lormand, E. (1996). "Nonphenomenal Consciousness," *Noûs*, 30: 242–61.
- Nagel, T. (1974). "What Is It Like to Be a Bat?" *The Philosophical Review*, vol. 82, 435–50.
- Nelkin, N. (1989). "Propositional Attitudes and Consciousness," *Philosophy and Phenomenological Research*, vol. 49, no. 3, 413–30.
- Nichols, S. and Stich, S. (2003). "How to Read Your Own Mind: A Cognitive Theory of Self-Consciousness," in Q. Smith and A. Jokic (eds), *Consciousness: New Philosophical Perspectives*. Oxford: Clarendon Press.
- Pitt, D. (2004). "The Phenomenology of Cognition, Or What Is It Like to Think that P?" *Philosophy and Phenomenological Research*, LXIX, no. 1.
- Pylyshyn, Z. (2001). *Seeing and Visualizing: It's Not What You Think*. Cambridge, MA: Bradford Books/MIT Press.
- Siegel, S. (2006). "Which Properties are Represented in Perception?" in T.S. Gendler and J. Hawthorne (eds), *Perceptual Experience*. Oxford: Oxford University Press.
- (2007). "How Can We Discover the Contents of Experience?" *The Southern Journal of Philosophy*, vol. xlv, 127–42.
- Siewert, C. (1998). *The Significance of Consciousness*. Princeton, NJ: Princeton University Press.