

Cognitive Phenomenology as the Basis of Unconscious Content

Uriah Kriegel Arizona

Forthcoming in *Cognitive Phenomenology*, eds. T. Bayne and M. Montague, Oxford UP

Abstract. Since the seventies, it has been customary to assume that intentionality is independent of consciousness. Recently, a number of philosophers have rejected this assumption, claiming that intentionality is closely tied to consciousness, inasmuch as non-conscious intentionality in some sense depends upon conscious intentionality. Within this alternative framework, the question arises of how to account for *unconscious* intentionality, and different authors have offered different accounts. A central goal of this paper is to argue for a broadly Dennettian, interpretivist account of unconscious intentionality. A second goal is to argue that an upshot of interpretivism is that all unconscious intentionality is ultimately grounded in a specific kind of cognitive phenomenology, namely, the phenomenology of conscious interpretive acts.

1. Introduction: Phenomenal Intentionality and the Question of Unconscious Intentionality

This paper's title pays tribute to Brian Loar's seminal paper "Phenomenal Intentionality as the Basis of Mental Content." Whereas mainstream analytic philosophy of mind has by and large assumed that the phenomenon of intentionality can be understood independently of the phenomenon of consciousness, Loar's paper (which was circulating since the late nineties) brought to the fore the possibility that all intentional content might be somehow grounded in a kind of distinctively conscious intentionality, which he called "phenomenal intentionality."

One way to precisify Loar's thesis is by appeal to the familiar distinction between derived and underived intentionality. Grice (1957) famously argued that the intentionality

of language is merely derivative, in that a linguistic symbol such $c^{\wedge}a^{\wedge}t$ is not in and of itself more suited for representing cats than dogs, and ultimately derives its content from speaker intentions, whereas the intentionality of the mental is non-derivative, in that mental states represent what they do in and of themselves. One way to interpret Loar's thesis is as claiming that in reality only phenomenally conscious states represent what they do in and of themselves, i.e. non-derivatively, while the intentionality of phenomenally *unconscious* states ultimately derives from phenomenal intentionality. (Note well: by "phenomenally unconscious" states I mean states that are unconscious in the phenomenal sense of "conscious"; I do not mean to refer, or even suggest the existence of, states that are unconscious but nonetheless have a phenomenal character.)

Independently of Loar exegesis, the emerging thesis is quite important. We may formulate it as follows:

(PI) The only non-derivative intentionality is phenomenal intentionality.

This thesis has been the cornerstone of an approach to intentionality which has been gathering momentum in recent years, in large part due to the work of Loar, as well as of such philosophers as Searle, Strawson, and Horgan, and which we may call the Phenomenal Intentionality Research Program (see Kriegel and Horgan forthcoming).¹ The central point of this program can be appreciated by contrast to the kind of approach to intentionality that sees it as altogether independent of consciousness.

Since work on intentionality started flourishing in the late nineteen-seventies, the dominant research program for understanding intentionality has been what we may call the Naturalist-Externalist Research Program (NERP). There are many ways to characterize this program; here is one back-of-the-envelope characterization: the source of intentionality is a certain natural relation that obtains between internal states of the brain and external states of the world; all other intentionality is ultimately based on this natural relation. Different theories within this research program offer different accounts of (a) the relevant natural relation and/or (b) the way it grounds all other intentionality. Prominent theories are associated with philosophers such as Dretske, Fodor, Harman, and Millikan.²

The Phenomenal Intentionality Research Program (PIRP) can be modeled on the same theoretical structure. Here the rough idea is that the source of all intentionality is the phenomenal character of conscious states, typically construed as a purely internal affair; all other intentionality occurs in virtue of appropriate relations to phenomenally conscious states. Again, different theories within PIRP would differ most centrally on (a) the kind of phenomenal character they take to constitute intentionality (and how it does so) and/or (b) the relation to phenomenal consciousness they take to underlie unconscious intentionality.³

Several important questions arise within the framework of PIRP. Perhaps the most central are the following three:

- Why should we believe the thesis that only phenomenal intentionality is non-derivative?
- What is the nature of phenomenal intentionality? In particular, what is it about phenomenal character that endows it with intentional content?
- What is the nature of non-phenomenal intentionality? More specifically, what is the relation a non-phenomenal item must bear to a phenomenally intentional state in order to exhibit intentionality?

What I want to do in the bulk of this paper is pursue this last question. There may arise for some readers legitimate worries about the viability and merit of PIRP, but I propose to set these worries aside here.⁴ The reason is partly substantive but mostly methodological. Substantively, it is simply my opinion, for which I argued elsewhere (Kriegel 2003a, 2007, forthcoming Ch.1) but will not here, that PIRP is in fact viable and merits pursuit. Methodologically, research programs in both philosophy and science rarely win converts by *a priori* demonstration of their viability and merit; rather, it is when the research program is actively pursued, and its pursuit is shown to be fruitful and productive, that it flourishes and attracts proponents and exponents. Actually pursuing PIRP is therefore central to making the case for its viability and merit.^{5,6}

Our central agenda is thus to seek a satisfactory account of unconscious intentionality within the PIRP framework. Call this the *unconscious intentionality question*:

(UIQ) For any phenomenally unconscious item x , such that x has intentional content, what makes it the case (i) that x has the intentional content it does and (ii) that x has an intentional content at all?

In §2, I propose an answer to UIQ, an answer that takes its cue from Dennett's (1971, 1981, 1987) interpretivist approach to intentionality. On the account I offer, what makes a phenomenally unconscious item have the intentional content it does, and an intentional content at all, is (very roughly) that it is profitably interpreted to have that content. For reasons we will see in §4, an upshot of this account is that all phenomenally unconscious intentionality derives ultimately from a certain type of cognitive phenomenology, a notion I will introduce and motivate in §3. The upshot, then, is that without cognitive phenomenology there would be no unconscious intentionality. I will close with more general remarks on the crucial role of cognitive phenomenology in the relationship between mind and world.

2. Interpretivism

(PI) implies that all unconscious intentionality derives from conscious intentionality.⁷ Interestingly, there is a NERP-ish approach of intentionality that takes *all* intentionality to be derived. The idea of this section is to adopt this approach but restrict it to unconscious intentionality (in an independently motivated way).

The approach I have in mind is Dennett's interpretivist "intentional stance" theory.⁸ On this view, there is a web of intentional concepts, such as of belief and desire, that we can use to produce a rough-and-ready interpretation of conspecifics and other creatures in real time and "on the go." When we use this web of intentional concepts, we take a stance toward our targets of interpretation that involves conceiving of them as intentional systems; this is the intentional stance. The intentional stance produces a kind

of interpretation that is a good enough approximation of the truth to make it useful, but not good enough that we can take it at face value. If we operated without constraints on the employment of time, energy, and other resources, we could produce a much more accurate theory of conspecifics' behavior, in principle a fully accurate one; but this other interpretation would deploy a web of neurophysiological and broadly physical concepts, not the web of intentional concepts we use in everyday life. In everyday life, the most cost-beneficial balance between resource expenditure and interpretative accuracy happens to be that provided by the intentional stance.

On this picture, it is not so much because certain intentional facts hold that certain interpretations become possible as that because those interpretations are possible that the intentional facts hold. Thus, for someone to believe that p just is for the best exercise of the intentional stance to assign to them the belief that p . As Dennett (1981: 72) puts it, "all there is to really and truly believing that p is being an intentional system for which p occurs as a belief in the best (most predictive) interpretation." On this interpretivist view, for an internal state to have intentional content C is for the state's subject to be best interpreted (within the framework of the intentional stance) as being in an internal state with C .

Interpreting someone to be in a certain intentional state is itself an intentional act, however: the act whose content is that someone is in some intentional state. Therefore, Dennett's interpretivism entails that the intentionality of each intentional state derives from the intentionality of some *other* intentional state, namely, the relevant interpretive state (whose own content presumably derives from that of a second-order interpretive act). It is thus an upshot of Dennett's view that all intentionality is derived (see Dennett 1988, 1990).

An immediate difficulty with this position is that it leads straightforwardly to infinite regress: an item x can acquire the content C only if there is an item y with content C^* , where $C^* = \langle x \text{ has } C \rangle$, but for y to have that content, there would have to be an item z with content C^{**} , $C^{**} = \langle y \text{ has } C^* \rangle$, and so on *ad infinitum*.⁹ The only way to stop the regress is to posit a class of privileged intentional states, such that (i) their intentionality does not derive from interpretation and (ii) they have interpretive acts as a subset.¹⁰

My suggestion is to designate phenomenally conscious intentional states as that privileged class, with phenomenally conscious interpretive acts as the requisite subset. Some independent story would have to be told about what endows phenomenally *conscious* intentional states with their intentional content; for that, see Kriegel forthcoming Chs.2,3. But once they do, and assuming some of them are phenomenally conscious acts of interpretation, we can account for the intentionality of phenomenally unconscious states as derived from theirs. To a first approximation, the resulting answer to UIQ would be this: what makes an unconscious item have the intentional content it does, and to have one at all, is that the best exercise of the intentional stance would produce an interpretation according to which that item has the intentional content it does, and has one at all.¹¹

Note that in the first instance, the relevant interpretation need not itself be phenomenally conscious. However, if it is not, then as noted above, it must have its own (interpretive) content in virtue of being interpreted by some second-order interpretation, and that second-order interpretation would itself be either conscious or interpreted by a third-order interpretive state, and so on. Ultimately, this regress of interpretations would have to end with a conscious interpretive act. In other words, the relevant interpretations need not be conscious, but they must be part of (potentially one-stepped) chains of interpretations that are capped by conscious interpretations.

We can simplify somewhat the formulation of this interpretivist account of unconscious intentionality by construing unconscious intentionality as the right sort of *response-dependent property* (Kriegel 2010a). These are properties characterized in terms of their disposition to elicit certain responses in certain respondents under certain conditions.¹² Since we are dealing here with the *best* exercise of the intentional stance, we can think of the relevant respondents, in the first instance, as *ideal interpreters*, that is, subjects who exercise the intentional stance perfectly under all conditions. The thesis is then that for any unconscious intentional state, there is *some* possible ideal interpreter who, under *some* conditions, produces an intentional interpretation of that state, and moreover does so *consciously*.¹³ If so, we may formulate interpretivism relatively economically as follows:

- (A) For any unconscious item x and intentional content C , x has C iff x is such as to elicit in ideal interpreters under ideal conditions a conscious intentional state with the content $\langle x$ has $C \rangle$.

Or perhaps more rigorously:

- (A1) For any unconscious item x and intentional content C , x has C iff there are a possible conscious item y , a possible ideal interpreter N , and possible conditions K , such that (i) y has the content $\langle x$ has $C \rangle$ and (ii) x is such as to elicit y in N under K .

I will now offer some clarifications (see also Kriegel 2010a) – in particular, concerning the nature of interpretation, ideal interpreters, and ideal conditions – and then suggest that this view is superior to the alternatives.

I start with clarification of the nature of interpretation. There are, in fact, several possible views on the nature of interpretation, and different versions of interpretivism will differ in adopting different views on that matter. In the literature on interpretation, three main theories seem to compete (Goldman 1989): the theory theory, simulation theory, and rationality theory.¹⁴ According to theory theory, intentional ascription is based on something like inference to the best explanation from the behavior (including verbal behavior) of the target of ascription, the kind of inference characteristic of theorizing in general. According to simulation theory, intentional ascription is based on the off-line tokening of the very intentional states ascribed, by way of simulation of the target of ascription's internal goings-on. According to rationality theory, intentional ascription is based first and foremost on the application of a principle of charity that casts the target of ascription as a rational agent whose intentional states are by and large justified. In addition to these three "pure" theories that assume all intentional ascription is carried out by a single mechanism, there are also "impure" theories that combine two or three of the above mechanisms. If the above are the *only* three interpretive mechanisms, this would mean that there are seven possible views of how interpretation works.¹⁵

For my part, although I cannot argue for this here, it seems to me that a simulation mechanism is most plausible for the ascription of phenomenal states, including phenomenally intentional states. So it would not be relevant as much for the ascription of phenomenally *unconscious* states, including unconscious intentional states. For those, my view is that interpretation involves the cooperation of a theory mechanism and a rationality mechanism. More specifically, ascription of unconscious intentional states proceeds in two phases: first, inference to the best explanation produces a number of possible ascriptions fully consistent with the behavioral explananda; then, the principle of charity is used to pick among them the single ascription that casts the target of interpretation in the most rational light.¹⁶ Thus my version of interpretivism appeals to a combination of theorization and rationalization as the elements of the relevant kind of interpretation.

It may be objected that this view cannot account for altogether non-mental intentionality, such as we find in language, traffic signs, representative art, etc., since it is unclear what the behavioral data for inference to the best explanation are supposed to be. The interpretivist has two options here, it seems to me. One is to insist on the theorization-cum-rationalization for non-mental intentionality, but offer different data for the relevant inferences to the best explanation. Another, perhaps more promising option is to ground non-mental intentionality in unconscious intentionality, thus offering a two-step derivation of non-mental intentionality from phenomenal intentionality. For example, an interpretivist could hold that non-mental intentionality derives from speaker intentions, as per familiar Gricean intention-based semantics. These speaker intentions are unconscious, and therefore derive their own content (according to interpretivism) from conscious interpretation. This would require the Gricean account to apply not only to linguistic representation, but indeed it has recently been argued that the Gricean account applies just as well to depictive representation (Abell 2005, Blumson 2006).¹⁷ On the emerging view, all intentional content derives ultimately from phenomenally intentional content, though it may derive from it in several stages. I am attracted to this view.

Having clarified the nature of interpretation, let us consider the nature of an ideal interpreter. I said that an ideal interpreter is one who exercises the intentional stance

perfectly under all conditions, but more can be said. For starters, we may construe an interpreter (in general) as any subject capable of entering interpretive states, that is, states of ascribing content to something. How to characterize what makes an interpreter ideal is not a simple matter. One option is virtue-epistemological: an ideal interpreter is a well-informed interpreter who exhibits (to the highest degree) all the epistemic virtues that bear on the production of interpretive states. Another option is teleofunctionalist: an ideal interpreter is a well-informed interpreter whose mechanisms responsible for the production of interpretive states always function exactly as they are supposed to. The best characterization for our present purposes, however, is to simply construe the ideal interpreter as an epistemic agent parachuted into the world (so to speak) with no knowledge of the facts we are appealing to her to fix but complete knowledge of all other facts (as well as perfect reasoning capacities). That is: an ideal interpreter is one who approaches an interpretive task with (i) complete knowledge of all the non-intentional facts and all the phenomenal-intentional facts, (ii) no knowledge of any other facts, (iii) the capacity to draw every valid deductive inference and every justified non-deductive inference, and (iv) the capacity to avoid drawing any invalid deductive inference and any unjustified non-deductive inference.¹⁸ This, if anyone, is an ideal interpreter.

Having now clarified the nature of interpretation and the ideal interpreter, let us next consider the nature of ideal conditions. Let us say that a condition K is ideal relative to a subject S and a task T just in case S can, under C, perform T the most competently that S can perform T. Given this construal of ideal conditions, and the above construal of the ideal interpretation, it would seem that any conditions are ideal relative to the ideal interpreter and the task of interpretation.¹⁹ However, it may be worthwhile to insist that the conditions relevant to the formulation of interpretivism are conditions of forced choice: the ideal interpreter must either assent to “x has C” or dissent from it – suspending judgment is not an option.

Note that an ideal interpreter, because ideal, would also (correctly) assign content to phenomenally intentional states.²⁰ So (A) holds not only of unconscious items but also of conscious items. However, the reason it holds is very different in each case. With *conscious* intentionality, the ideal interpreter *tracks* intentionality. With *unconscious* intentionality, s/he *constitutes* intentionality. There is something akin to a Euthyphro

Dilemma here: does the ideal interpreter assign content C to item x because x has C or does x have C because the ideal interpreter assigns to it C ? With conscious intentionality it is the former, with unconscious intentionality the latter. Crispin Wright (1992) claims that such a dilemma applies to all biconditionals of this form, and when the biconditional is supposed to capture a genuinely response-dependent property, it must hold *a priori*. If so, we should probably modify (A1) as follows:

- (A2) Epistemically necessarily, for any unconscious item x and intentional content C , x has C iff there are a possible conscious item y , a possible ideal interpreter N , and possible conditions K , such that (i) y has the content $\langle x \text{ has } C \rangle$ and (ii) x is such as to elicit y in N under K .

In this formulation, I use epistemic necessity as a formally well-behaved gloss on aprioricity. Roughly, p is epistemically necessary iff p is true in every centered world, i.e., in every possible world considered as actual.²¹

It is also important to note that (A)-(A2) do not require in any way the existence of an actual interpreter, let alone an ideal one. They insist that an unconscious item must have the *disposition* to elicit the right interpretation in the right interpreters, but not that the disposition must be *manifested*. The non-existence of ideal interpreters only means that the relevant disposition cannot be manifested, not that it is not present.

There may be cases where there is no single best interpretation of some item. In those cases, one ideal interpreter would ascribe to an item content C_1 and another would ascribe to it content C_2 . In such circumstances, (A) entails that the item's content is indeterminate as between C_1 and C_2 . But this kind of content indeterminacy should be extremely infrequent, and to that extent harmless – it is not the kind of corrosive indeterminacy threatened by inscrutability worries. For there is no reason to suppose that in standard cases there is more than one best interpretation. Crucially, since the ideal interpreter knows all the phenomenal-intentional facts, it knows that ordinarily I consciously think of rabbits rather than undetached rabbit parts, and would therefore interpret my relevant *unconscious* states as about rabbits not undetached rabbit part. Such

an interpretation would make better sense of the cognitive, inferential interactions among my various intentional states, some of which are conscious and some unconscious.

Our interpretivism about unconscious intentionality would bring in its train all the hallmarks of the more standard, global type of interpretivism we find in Dennett, Davidson, and others. For example, interpretivism often goes hand in hand with *holism*, the view that contents cannot be assigned to individual items but must be assigned to whole groups at once. This is because creatures' behavior admits of several coherent interpretations in which changes in the assignment of content to one item are compensated for by changes in the assignment of content to another item. For example, a student who shows up in the wrong room on the day of the exam may be interpreted either as (i) believing that the exam is in (what is in fact) the wrong room and desiring to take the exam or (ii) believing that the exam is in (what is in fact) the right room and desiring not to take the exam. This kind of holism certainly characterizes the interpretation of behavior caused by unconscious mental states – indeed, we may suppose that our student's beliefs and desires are unconscious. Thus interpretivism about unconscious intentionality, as captured in (A2), results in holism about unconscious content.

A related doctrine often associated with interpretivism is that massive error and irrationality are a priori impossible. This is because when several coherent interpretations are available, competent selection among them appeals to a principle of charity. For example, if your interlocutor points at your new laptop and says “congratulations on your new capsicum,” it is possible to interpret her as either (i) desiring to congratulate you on a new pepper and believing that the word “capsicum” means pepper or (ii) desiring to congratulate you on a new laptop and believing that the word “capsicum” means laptop. A competent interpreter would always opt for the second interpretation, even though both accommodate the data equally well, because (presumably) it is the more charitable interpretation. The combination of holism and the employment of the principle of charity then tends to undermine the ascription of erroneous and/or irrational contents to other items in the relevant web of unconscious items, with the result that sweeping error and irrationality are ruled out a priori.

Another doctrine often associated with interpretivism is a kind of irrealism or instrumentalism about intentionality. According to such instrumentalism, intentionality is a useful fiction: the practice intentional ascription is useful for the conduct of everyday life, but does not strictly speaking issue in true descriptions of what actually goes on in the world. At the same time, while ascription of unconscious intentionality is certainly useful, and moreover its existence cannot be appreciated independently of its usefulness according to interpretivism, it does not follow that it is a *fiction*. In fact, a common form argument reasons from usefulness to truth via inference to the best explanation (Fodor and Lepore 1993). The issues raised by this question are too involved to address here, but let me go on the record expressing my preference for a non-instrumentalist reading of (A2). After all, the response-dependent property is fully integrated in the natural order, is publicly accessible, etc.²²

This, then, is the interpretivism about unconscious intentionality that I propose. What motivates the view to me is that it is vastly preferable to the alternative accounts of unconscious intentionality one can find in the literature. I argue for this in detail in Kriegel forthcoming Ch.4, where I consider three alternative accounts, which I call *potentialism*, *inferentialism*, and *eliminativism*. According to potentialism (Searle 1991, 1992), an unconscious state has the intentional content it does in virtue of potentially having the phenomenal character it potentially does, and has intentional content at all in virtue of potentially having phenomenal character at all. According to inferentialism (Loar 2003, Horgan and Graham 2009), what makes an unconscious state have intentional content is that it is inferentially or functionally integrated into a cognitive system in which some intentional states are phenomenally conscious; what makes it have the specific intentional content it does, rather than another, is that it plays the specific inferential or functional role it does, and not another, within that cognitive system. According to eliminativism (Strawson 2008, Georgalis 2006), meanwhile, there simply is no unconscious intentionality. The considerations bearing against these accounts are many and varied, but strikingly, all three appear to get wrong the extension of unconscious intentionality. They leave us with less intentionality than we are antecedently inclined to think there is. Potentialism cannot account for the intentionality of unconscious states that are not even potentially conscious, such as visual states in the

dorsal stream (Davies 1995, Horgan and Kriegel 2008). Inferentialism cannot account for the intentionality of inferentially insulated unconscious states, as well as of zombie states. And eliminativism cannot account for *any* antecedently plausible instance of unconscious intentionality. For details, see Kriegel forthcoming Ch.4.

My claim is that interpretivism returns the right results in all cases in which it seems antecedently that we should count a certain unconscious state as intentional. Importantly, there is something principled in the way interpretivism manages this. The reason is that what makes it desirable to count something as in the extension of intentionality is not accidentally related to what makes it the case that interpretivism returns the result that it is a case of intentionality. There are two kinds of fact that make it desirable to count a given state as intentional: (i) that it is so treated by cognitive science and (ii) that it is so treated by folk psychology. Certainly in the former case, but probably also in the latter, what makes it desirable to count the state as intentional is that it is explanatorily beneficial to do so. But whenever it is explanatorily beneficial to treat a state as intentional, an ideal interpreter would do so, and therefore interpretivism would ratify the state's status as intentional. Thus it is no accident that interpretivism returns the independently desirable results in the above cases and would probably do so in other cases. The only way it could fail to do so is if there were unconscious states to which we have independent reasons to ascribe intentionality but for which it is explanatorily useless to do. In all likelihood, however, there could be no such states: unlike phenomenally intentional states, we have no first-person acquaintance with unconscious intentional states; our only reason to believe in them is that doing so is theoretically or explanatorily profitable from the third-person perspective.

In the assessment of scientific theories, it is common to distinguish empirical and theoretical virtues: the empirical virtues concern accounting for the data, the theoretical virtues concern parsimony, simplicity, unity, clarity, elegance, conservatism, etc. It is natural to think that a similar distinction applies to philosophical theories, where the independently desirable results play the role of data, such that returning those results is the empirical virtue (or "quasi-empirical," if we prefer).²³ It is clear from the above discussion that interpretivism is more (quasi-)empirically virtuous than its competition.²⁴

It is, in effect, massively *confirmed*, whereas alternative accounts are often enough *disconfirmed*.

3. Cognitive Phenomenology

The most phenomenologically impressive experiences we undergo are visual experiences and somatic experiences: color perception and pain/pleasure, in particular, are often phenomenologically overwhelming. But there is a variety of milder phenomenologies we often experience. For starters, other perceptual phenomenologies, such as olfactory and proprioceptive phenomenology, can be quite subtle and unimposing, as can the phenomenology of sufficiently mild background moods.

Indeed, some phenomenologies may be so unimpressive that it becomes controversial whether they exist. There are many examples of this, but the one I will focus on here is what is often referred to as *cognitive phenomenology*. The idea is that there is a distinctive phenomenology involved in having an occurrent thought, which moreover is not just the phenomenology of the imagery accompanying the thought, but is proprietary to occurrent thoughts.²⁵ To say that this phenomenology is proprietary, or *sui generis*, is to say that it does not reduce to some other, already familiar (perceptual or somatic, in particular) phenomenology.

It is, of course, controversial whether such proprietary cognitive phenomenology exists, but several authors have recently argued that it does (Goldman 1993, Strawson 1994, Peacocke 1998, Siewert 1998, Horgan and Tienson 2002, Kriegel 2003b, Pitt 2004, Klausen 2008, Tennant 2009).²⁶ The view that it exists admits of several grades (Kriegel 2003b). The strongest claim would be that every type of occurrent thought has its own cognitive phenomenology, such that the cognitive-phenomenal character of believing that p is different from that of believing that q (whenever $p \neq q$). A weaker claim would be that cognitive phenomenology varies with attitude but not with content, such that believing that p is phenomenally different from desiring that p but not from believing that q .²⁷ The weakest claim is that there is a single cognitive-phenomenal character shared by all conscious cognitive states. Proponents of PIRP typically assert the strongest of these claims. And so they should: the more phenomenology there is, the more phenomenal

intentionality there is, and the more plausible it becomes that all intentionality derives from phenomenal intentionality.²⁸

In the literature, there are three main argumentative strategies that have been wielded by way of defending the existence of cognitive phenomenology: arguments from phenomenal contrast, arguments from phenomenological overwhelm, and arguments from first-person knowability.

In arguments from phenomenal contrast, two overall conscious episodes of a subject are presented, such that (i) it is intuitively clear that there is an overall phenomenal difference between the two, and (ii) the best explanation of this difference is that one of the two episodes exhibits a cognitive phenomenology that the other does not. Perhaps the best known instance of this argumentative strategy is Strawson's (1994) contrast of the overall experiences of a French speaker and a non-French-speaker when they listen to the news in French. Strawson argues that there is a phenomenal difference between their overall conscious episodes, and that the difference is best accounted for in terms of an element of understanding-experience present only in the French speaker's stream of consciousness.²⁹

In arguments from phenomenological overwhelm, the contrast between two episodes is replaced with a single episode, clearly cognitive in nature, whose phenomenal character is supposed to be overwhelming in the way that approximates the phenomenological overwhelm involved in visual and somatic experiences and due to which they are unquestionable instances of phenomenality. Perhaps the best known instance of this argument is based on the sudden and acute onset of an experience of grasping something. Consider the following passage (Mangan 2001, Chudnoff forthcoming):

A newspaper is better than a magazine. A seashore is a better place than the street. At first it is better to run than to walk. You may have to try several times. It takes some skill but it is easy to learn. Even young children can enjoy it. Once successful, complications are minimal. Birds seldom get too close. Rain, however, soaks in very fast. Too many people doing the same thing can also cause problems. One needs lots of room. If there are no complications it can be very peaceful. A rock will serve as an anchor. If things break loose from it, however, you will not get a second chance.

This passage elicits mostly puzzlement, until one is told that it is about *kites*. Once informed, one typically undergoes the acute onset of the grasp experience of which I speak.³⁰

In arguments from first-person knowability, it is pointed out that the knowledge we have of some of our occurrent cognitive states is the kind of knowledge we have only of our phenomenally conscious states, from which it is inferred that these cognitive states must be phenomenally conscious.³¹ Thus, Pitt (2004) argues that we have an immediate, non-inferential knowledge of some cognitive states (as well as their contents), and that only phenomenally conscious states are knowable in this way. He concludes that there is a phenomenal character proper to cognitive states (and their contents).³²

Opponents have tended to be unmoved by these considerations, claiming that there are better explanations of phenomenal contrast, phenomenal overwhelm, and first-person knowability than the supposition of cognitive phenomenology. Frustratingly, the disagreement often seems to devolve into a seemingly verbal matter rather quickly. Both sides agree that cognitive states can exhibit some special feature F, but disagree on whether F is a phenomenal feature, and it is unclear in what way the issue goes beyond whether one is willing to apply the term “phenomenal” to F.

Under these conditions, progress would require first and foremost that a characterization of phenomenality be devised which is antecedently neutral and does not prejudge the question of the existence of cognitive phenomenology. In Kriegel 2009 Ch.1, I have suggested that phenomenal consciousness is best characterized as the “explanatory-gap-able property.” More specifically, I suggested that we fix the reference of “phenomenality” with the following rigidified definite description: “The property P, such that, in the actual world, there is (the appearance of) an explanatory gap between physical properties and P.”³³ Any species of the genus P would then qualify as a phenomenal property. (For fuller development of this criterion of phenomenality, see Kriegel Ms.)

Although I cannot expand on this here, it seems to me that an argument could be erected for cognitive phenomenology which would work with this characterization. This would involve conceiving a world in which subjects have no perceptual, somatic, or

emotional phenomenology and claiming that the worry captured by the explanatory gap nonetheless appropriately arises for this world. There are conditions involving congenital absence of perceptual phenomenology, such as blindness and deafness, as well as conditions involving congenital absence of somatic phenomenology (e.g., congenital analgesia) and of emotional phenomenology (e.g., perhaps, in certain forms of autism).³⁴ We can certainly envisage a person suffering from all these conditions at once (Kriegel 2003b, Ms), and can therefore conceive of a world where everybody does. The proponent of cognitive phenomenology would then make the following substantive claim: it is rationally appropriate to feel the pull of the explanatory gap for that world. This claim suggests that the explanatory-gap-able property is instantiated in that world, and therefore that phenomenality is instantiated. Since perceptual, somatic, and emotional phenomenologies are *not* instantiated, it follows that there must be a non-perceptual, non-somatic, and non-emotional phenomenality. One could then argue that the best candidate for this non-perceptual, non-somatic, and non-emotional phenomenality is cognitive phenomenality. This would be again not a demonstrative argument, but an argument from inference to the best explanation. Its advantage over the previous kinds of argument is that it would not be as susceptible to devolution into verbal dispute.

I develop this sort of argument in more detail elsewhere (Kriegel Ms). Here I will *assume* the existence of (proprietary or *sui generis*) cognitive phenomenology, and will concern myself rather with its relation to unconscious intentionality. I want to argue that unconscious intentionality is grounded ultimately in a certain type of cognitive phenomenology, namely, the cognitive phenomenology of conscious interpretation.

4. Interpretivism and Cognitive Phenomenology

In this final section, I argue that an upshot of interpretivism is the proposition sloganeered in the title: cognitive phenomenology is the basis of unconscious content. The argument I will present is this:

- 1) Interpretivism entails that all unconscious content is ultimately grounded in conscious interpretation;

- 2) The content of all conscious interpretation is grounded in the phenomenal character of conscious interpretation;
- 3) The phenomenal character of all conscious interpretation is a kind of cognitive phenomenology; therefore,
- 4) Interpretivism entails that all unconscious content is ultimately grounded in a kind of cognitive phenomenology.

I have already defended the first premise – that was the mandate of §2. This section defends the second and third premises.

Within NERP, there is no reason to believe the second premise. But the premise is an essential – perhaps *the* essential – tenet of PIRP. The focus of this paper has been on the nature of *unconscious* intentionality within PIRP. The other half (if you will) of a PIRP-ish theory of intentionality concerns the theory of *conscious* intentionality. There are several options here too, but they all have in common that the intentionality of phenomenally conscious states is grounded in those states' phenomenal character.

It is not at all clear how the relevant grounding is to be cashed out, though it is natural to appeal to a certain relation of metaphysical dependence. The relation is surely anti-reflexive and transitive,³⁵ but it could be legitimately construed either as *asymmetric* or as *anti-symmetric*. The anti-symmetric interpretation would entail that the phenomenal character of a phenomenally intentional state does not depend on the state's intentional content, whereas the asymmetric interpretation would not have this entailment – it would be compatible with both one-way and two-way dependence. One could make this difference explicit by subscripting “metaphysical dependence.” Another device would be to construe metaphysical dependence as asymmetric and then explicitly distinguish two grades of the grounding thesis: a strong conjunctive thesis claiming that the intentional is grounded in the phenomenal and the phenomenal is not grounded in the intentional, and a weak thesis consisting merely of the first conjunct of the strong thesis. An intriguing version of the this weaker thesis would be an *identity* thesis: the intentional property and the phenomenal property are one and the same. In this thesis, it is clear that there is no primacy of the phenomenal over the intentional within phenomenal intentionality. If one hears “grounding” as implying primacy, this weaker thesis would not cast the intentional

as grounded in the phenomenal. But for present purposes, let us hear “grounding” as uncommitted on the primacy of the phenomenal to the intentional, and committed only regarding this: there is no primacy of the intentional to the phenomenal (as there is in standard intentionalist theories of phenomenal consciousness). With this stipulation, it is clearly a central tenet of PIRP that the intentional content of phenomenally conscious states is grounded in the phenomenal character of those states.

Naturally, this applies also to the intentionality of phenomenally conscious interpretive acts. Such acts, being phenomenally conscious, have phenomenal character. Moreover, it is plausible, though not uncontested, that there is some phenomenal commonality among them. Just as there is a phenomenal commonality among all gustatory experiences as of bitter things, which is the phenomenology of bitterness, and a subtler phenomenal commonality among all conscious thoughts about the number pi, which we may call the pi phenomenology, so probably there is a phenomenal commonality among all conscious acts of intentional ascription, which we may call *interpretive phenomenology*. This is not trivial: every phenomenally conscious state that occurs on a Wednesday has a phenomenal character, but there is no Wednesday phenomenology, because there is no phenomenal commonality among Wednesday experiences. But it appears to be the case that whenever phenomenally conscious states have an intentional commonality they also have a phenomenal commonality (though, again, this is not beyond rational doubt).³⁶ Thus the fact that there is an intentional commonality among conscious interpretive acts, in that they all ascribe an intentional property to something, suggests that there is a phenomenal commonality among them as well.³⁷ That is, it suggests that there is such a thing as an interpretive phenomenology. What the proponent of PIRP would typically claim is that this interpretive phenomenology grounds the intentional content of phenomenally conscious interpretive acts. She would, in other words, assert the second premise of the argument with which I opened this section.³⁸

It is not easy to say anything particularly illuminating about the nature of interpretive phenomenology. An involved phenomenological analysis may excavate interesting aspects of this phenomenology, but short of fully engaging in such analysis, there is not much to do beyond point out the intentional commonality interpretive

phenomenology grounds: what it is like to interpret someone as believing that it is sunny is just the someone-believes-that-it-is-sunny what-it-is-like.

Even without a comprehensive analysis of interpretive phenomenology, a case can probably be made that all interpretive phenomenology is cognitive. There are three main views about which kinds of mental state are phenomenally conscious interpretive acts. One view is that all conscious interpretive acts are beliefs, or judgments, or thoughts. On this view, to interpret someone as believing that it is sunny is to believe/judge/think that they believe that it is sunny. A second view is that some interpretive acts can be perceptual: one can *see* that someone believes that it is sunny; at the very least, *some* intentional states are such that one can just *see*, or otherwise *perceive*, that someone is in them. A third view is that interpretation involves rather some kind of *sui generis* attitude, the attitude of interpreting that *p*, e.g. that someone believes that it is sunny.

The first, “cognitive” view seems to me the most plausible. The third, “*sui generis*” view seems unmotivated: it is unclear what explanatory work cannot be done by any familiar attitude that a new kind of attitude should be posited. Certainly the burden of argument should be on those who wish to posit the unfamiliar “interpretive attitude.” As for the second, “mixed” view, it would only be well motivated if there were clear examples of perceptual interpretive acts. I think it is fairly clear that the case of interpreting someone to believe that it is sunny is not such an example – it is difficult to envisage in what way a person could perceive another to believe that it is sunny.

There may be other, more tempting cases, of course. Thus, perhaps one can just *see* that one’s friend is smelling something funny, say by seeing a particular facial contortion of one’s friend. In this case, one’s visual experience ascribes to one’s friend an intentional state, namely, an olfactory experience as of a foul odor.

If the first view is the right one, and interpretive acts are always cognitive, it would seem to follow quite straightforwardly that interpretive phenomenology is a variety of cognitive phenomenology, since all phenomenally conscious interpretive acts are construed as cognitive states (beliefs/judgments/thoughts). More interestingly, plausibly this is the case even if the second, “mixed” view is the right one. For it is plausible that the phenomenology of any perceptual interpretive acts is a composite of

sensory and cognitive phenomenologies. Observe that our example of such a perceptual interpretive act involves seeing that a friend is smelling something funny *by* seeing a facial contortion. Arguably, the phenomenal character of seeing the facial contortion exhausts the purely sensory component of the overall experience's phenomenal character. What one sees *by* seeing the facial contortion, the phenomenal component that goes beyond seeing the contortion, is a kind of non-sensory phenomenology. Moreover, it is hard to see what kind of phenomenology such non-sensory phenomenology could be if not cognitive phenomenology. It is certainly not a somatic or emotive phenomenology. Nor does it seem to be a conative or agentive phenomenology. This leaves cognitive phenomenology as the only natural candidate for understanding this non-sensory phenomenology involved in perceiving someone to smell something funny.

This treatment of such perception requires a conception of perceptual phenomenology as potentially comprised of two components, a sensory-phenomenology component and a cognitive-phenomenology component.³⁹ It is no part of this conception that the two components must be neatly factorizable. It may well be that perceptual phenomenology consists in a phenomenal fusion of the two components that does not allow for neat separation of their individual contributions. Nonetheless, to the extent that the phenomenal character of a perceptual act comprises this cognitive element or contribution, it is a kind of cognitive phenomenology. My contention is that perceptual acts lacking this kind of cognitive-phenomenal component, such as perceptual experiences of facial contortions, do not have interpretive phenomenology. The only mental acts that exhibit interpretive phenomenology are acts with cognitive phenomenology, and moreover, it is in virtue of their cognitive phenomenology that they exhibit interpretive phenomenology. In other words, their interpretive phenomenology is a kind of cognitive phenomenology.

It may be worthwhile to distinguish the emerging picture of perceptual interpretive acts from a certain historical conception of perception. According to this historical view, perceptual experiences are composites of sensation and belief: to perceive something is to (roughly simultaneously) sense something and believe something. The view suggested here is different. It does not posit two numerically distinct states, a sensation and a belief, but rather allows a perception to be a single unitary state. At the

same time, it has some continuity with the historical view, inasmuch as it finds structure in the perceptual phenomenology of that unitary state. This structure casts perceptual phenomenology as a composite of two kinds of phenomenology, sensation-like and belief-like.⁴⁰ In any event, the emerging picture portrays perceptual interpretive acts as exhibiting a composite phenomenology comprising a sensory component and a cognitive component, with their interpretive phenomenology being part of the cognitive component, that is, being a form of cognitive phenomenology.

Beyond its natural appeal, this picture is further supported by the fact that our example of a perceptual interpretive act is one of seeing *that*, therefore one of a *propositional attitude*, though a perceptual propositional attitude. As Klausen (2008) argues, the phenomenal character propositional attitudes have *qua* propositional attitudes is always a kind of cognitive phenomenology, and this applies to perceiving-that: insofar as one's perceiving is propositional, one's experience's phenomenal character is cognitive.⁴¹

I conclude that both on the first view of conscious interpretive acts, as always purely cognitive, and on the most plausible reading of the second, "mixed" view, according to which some interpretive acts are perceptual, interpretive phenomenology is a kind of cognitive phenomenology. This is the third premise of the argument at the opening of this section.

This completes my preliminary case for the reasoning that grounds unconscious intentionality, via conscious interpretation and interpretive phenomenology, in cognitive phenomenology. I do not pretend to have offered a thorough defense of this argument here. Rather, my aim has been to sketch the case for each of its premises. My hope, however, is that this section has made it *prima facie* highly plausible that an upshot of interpretivism about unconscious intentionality is that unconscious intentionality is ultimately grounded in cognitive phenomenology.

5. Conclusion

Cognitive phenomenology may be at the basis not only of unconscious intentionality, but also of conscious intentionality. For another question within PIRP concerns the

phenomenal signature of intentionality: what phenomenal feature, if any, transforms phenomenal *character* into phenomenal *intentionality*. On some views, that feature may turn out to be a kind of cognitive phenomenology as well (though presumably a different kind from interpretive phenomenology). One such view is Strawson's (2008): there is a subtle phenomenal feature whereby the mind *takes* something to be thus-and-so, and this *phenomenal taking* is a kind of cognitive phenomenology present in every conscious experience endowed with intentionality.^{42,43} Personally, I am not persuaded that phenomenal taking is this key feature; but someone who were, or who identified some other cognitive phenomenology as the key to phenomenal intentionality, and who also adopted interpretivism about unconscious intentionality, could claim cognitive phenomenology as the basis of both phenomenal and non-phenomenal intentionality, hence as grounding all intentionality.

I have not argued for this stronger claim, though I am somewhat sympathetic to it (see Kriegel forthcoming Ch.5). My main aim here has been to present the interpretivist approach to unconscious intentionality within PIRP, indicate what the case for it is, and argue that it ultimately grounds unconscious intentionality in a kind of cognitive phenomenology, since the intentional content of conscious interpretation is grounded in interpretive phenomenology, which is a kind of cognitive phenomenology. If all this is right, cognitive the phenomenology is the basis of unconscious content.⁴⁴

References

- Abell, C. 2005. "Pictorial Implicature." *Journal of Aesthetics and Art Criticism* 63: 55-66.
- Blumson, B. 2006. *Resemblance and Representation*. PhD Dissertation, Australian National University.
- Bourget, D, Forthcoming. "Consciousness is Underived Intentionality." *Noûs*.
- Brentano, F. 1874. *Psychology from Empirical Standpoint*. Edited by O. Kraus. English edition L. L. McAlister. Translated by A. C. Rancurello, D. B. Terrell, and L. L. McAlister. London: Routledge and Kegan Paul, 1973.
- Chalmers, D.J. Forthcoming. "The Nature of Epistemic Space." In A. Egan and B. Weatherson (eds.), *Epistemic Modality*. Oxford and New York: Oxford University Press.

- Child, W. 1994. *Causality, Interpretation, and the Mind*. Oxford: Oxford UP.
- Chudnoff, E. Forthcoming. "Intellectual Gestalts." In T. Horgan and U. Kriegel (eds.), *Phenomenal Intentionality: New Essays*. Oxford and New York: Oxford University Press.
- Davies, M. 1995. "Consciousness and the Varieties of Aboutness." In C. Macdonald and G. Macdonald (eds.), *Connectionism: Debates on Psychological Explanation* (Vol. II). Oxford: Blackwell.
- Dennett, D.C. 1971. "Intentional Systems." *Journal of Philosophy* 68: 87-106.
- Dennett, D.C. 1981. "True Believers." In A.F. Heath (ed.), *Scientific Explanation*. Oxford: Oxford UP. Reprinted in Dennett 1987.
- Dennett, D.C. 1987. *The Intentional Stance*. Cambridge MA: MIT Press.
- Dennett, D.C. 1988. "Evolution, Error, and Intentionality." In Y. Wilks and D. Partridge (eds.), *Sourcebook on the Foundations of Artificial Intelligence*. Albuquerque NM: New Mexico UP 1988.
- Dennett, D.C. 1990. "The Myth of Original Intentionality." In K.A. Mohyeldin Said, W.H.
- Dretske, F.I. 1980. "The Intentionality of Cognitive States." *Midwest Studies in Philosophy* 5: 281-294.
- Dretske, F.I. 1981. *Knowledge and the Flow of Information*. Oxford: Clarendon.
- Dretske, F.I. 1988. *Explaining Behavior*. Cambridge MA: MIT Press.
- Fodor, J.A. 1990. *A Theory of Content and Other Essays*. Cambridge MA: MIT Press.
- Fodor, J.A. and E. Lepore 1993. "Is Intentional Ascription Intrinsically Normative?" In Dahlbom (ed.), *Dennett and His Critics*, Oxford: Blackwell.
- Georgalis, N. 2006. *The Primacy of the Subjective*. Cambridge MA: MIT Press.
- Goldman, A.I. 1989. "Interpretation Psychologized." *Mind and Language* 4: 161-185.
- Goldman, A. 1993. "The Psychology of Folk Psychology." *Behavioral and Brain Sciences* 16: 15-28.
- Grandin, T. 1996. *Thinking in Pictures*. New York: Vintage.
- Grice, H. P. 1957. "Meaning." *Philosophical Review* 66: 377-88.
- Harman, G. 1987. "(Non-Solipsistic) Conceptual Role Semantics." In E. Lepore (ed.), *New Directions in Semantics*. London: Academic Press.
- Horgan, T. and G. Graham 2009. "Phenomenal Intentionality and Content Determinacy." In R. Schantz (ed.), *Prospects for Meaning*. Amsterdam: de Gruyter.
- Horgan, T. and U. Kriegel 2008. "Phenomenal Intentionality Meets the Extended Mind." *Monist* 91.
- Horgan, T. and J. Tienson 2002. "The Intentionality of Phenomenology and the Phenomenology of Intentionality." In D. J. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford and New York: Oxford UP.
- Horgan, T., J. Tienson, and G. Graham 2004. "Phenomenal Intentionality and the Brain in a Vat." In R. Schantz (ed.), *The Externalist Challenge: New Studies on Cognition and Intentionality*. Amsterdam: de Gruyter.

- Johnston, M. 1989. "Dispositional Theories of Value." *Proceedings of Aristotelian Society* 63: 139-174.
- Klausen, S.H. 2008. "The Phenomenology of Propositional Attitudes." *Phenomenology and the Cognitive Sciences* 7: 445-462.
- Kriegel, U. 2003a. "Is Intentionality Dependent upon Consciousness?" *Philosophical Studies* 116: 271-307.
- Kriegel, U. 2003b. "Consciousness as Sensory Quality and as Implicit Self-Awareness." *Phenomenology and the Cognitive Sciences* 2 (2003): 1-26.
- Kriegel, U. 2007. "Intentional Inexistence and Phenomenal Intentionality." *Philosophical Perspectives* 21: 307-340.
- Kriegel, U. 2009. *Subjective Consciousness: A Self-Representational Theory*. Oxford: Oxford UP.
- Kriegel, U. 2010a. "Interpretation: Its Scope and Limits." In A. Hazlett (ed.), *New Waves in Metaphysics*. London: Palgrave-Macmillan.
- Kriegel, U. 2010b. "Intentionality and Normativity." *Philosophical Issues* 20.
- Kriegel, U. Forthcoming. *The Sources of Intentionality*. Oxford and New York: Oxford University Press.
- Kriegel, U. Ms. "Cognitive Phenomenology: From Analysis to Argument."
- Kriegel, U. and T. Horgan forthcoming. "The Phenomenal Intentionality Research Program." In T. Horgan and U. Kriegel (eds.), *Phenomenal Intentionality: New Essays*. Oxford and New York: Oxford University Press.
- Loar, B. 1995. "Reference from the First-Person Perspective." *Philosophical Issues* 6: 53-72.
- Loar, B. 2003. "Phenomenal Intentionality as the Basis for Mental Content." In M. Hahn and B. Ramberg (eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. Cambridge MA: MIT Press.
- Lycan, W.G. 2008. "Phenomenal Intentionalities." *American Philosophical Quarterly* 45: 233-252.
- McDowell, J. 1985. "Values and Secondary Qualities." In T. Honderich (ed.), *Morality and Objectivity: A Tribute to John Mackie*. London: Routledge and Kegan Paul.
- McGinn, C. 1988. "Consciousness and Content." In *Proceedings of the British Academy* 76: 219-239. Reprinted in N. J. Block, O. Flanagan, and G. Güzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*. Cambridge MA: MIT Press, 1997.
- Mangan, B. 2001. "Sensation's Ghost: The Non-Sensory Fringe of Consciousness." *Psyche* 7.
- Millikan, R.G. 1984. *Language, Thought, and Other Biological Categories*. Cambridge MA: MIT Press.
- Moore, G.E. 1953. "Propositions." In his *Some Main Problems of Philosophy*. Oxford: Routledge.
- Peacocke, C. 1998. "Conscious Attitudes, Attention, and Self-Knowledge." In C. Wright, B.C. Smith, and C. Macdonald (eds.), *Knowing Our Own Minds*. Oxford: Oxford University Press.

- Pitt, D. 2004. “The Phenomenology of Cognition; or *What Is It Like to Think that P?*”, *Philosophy and Phenomenological Research* 69: 1-36.
- Searle, J. R. 1991. “Consciousness, Unconsciousness, and Intentionality.” *Philosophical Issues* 1: 45-66.
- Searle, J.R. 1992. *The Rediscovery of Mind*. Cambridge MA: MIT Press.
- Siewert, C. 1998. *The Significance of Consciousness*. Princeton NJ: Princeton UP.
- Siewert, C. This volume. “Phenomenal Thought.”
- Strawson, G. 1994. *Mental Reality*. Cambridge MA: MIT Press.
- Strawson, G. 2005. “Intentionality and Experience: Terminological Preliminaries.” In D.W. Smith and A. Thomasson (eds.), *Phenomenology and Philosophy of Mind*. Oxford and New York: Oxford UP.
- Strawson, G. 2008. “Real Intentionality 3: Why Intentionality Entails Consciousness.” In G. Strawson, *Real Materialism and Other Essays*. Oxford: OUP.
- Tennant, N. 2009. “Cognitive Phenomenology, Semantic Qualia, and Luminous Knowledge.” In P. Greenough and D. Pritchard (eds.), *Williamson on Knowledge*. Oxford: Oxford UP.
- Wright, C. 1992. *Truth and Objectivity*. Cambridge MA: Harvard UP.

¹ See Searle 1991, 1992, Loar 1995, 2003, Strawson 1994, 2005, 2008, Horgan and Tienson 2002, Horgan and Graham 2009. Other defenses of (PI) include McGinn 1988, Horst 1996, Kriegel 2003a, 2007, forthcoming, Georgalis 2006, Bourget forthcoming, and as far as I can tell Brentano 1874 and Husserl 1900.

² See Dretske 1980, 1981, 1988, Fodor 1990, Millikan 1984, Harman 1987. For example, Dretske (1988) combines a teleo-informational account of the all-important natural relation and an intention-based account of the way it grounds other intentionality. The view is that internal states of certain systems are intentional in virtue of being recruited to perform a certain function within the system due to the informational connections they bear to external conditions, and anything intentional that is not intentional for this reason is intentional because we use it with a certain intention (the intention having its intentionality from the relevant natural relation).

³ Later on I will discuss in passing some of the suggestions that have floated for understanding how phenomenal character grounds intentionality in phenomenal intentionality. For a more thorough discussion, see Kriegel forthcoming Ch.3.

⁴ Proponents of NERP are typically mystified by the notion that phenomenal character can blossom into intentionality: they allow that there may be a sense in which a phenomenal character can “fit” an external object, but suspect that it would also fit any qualitatively indistinguishable object. However, a parallel worry about the sources of intentionality also animates many proponents of PIRP, who are equally mystified as to how external connections can blossom into intentionality: such relations can perhaps “connect” an internal item to an external object, but presumably they would connect it to any perfectly collocated external item. (For example, they would connect the internal item just as well to the fusion of the external object’s undetached parts. For this consideration, see Loar 1995, Horgan and Tienson 2002, Georgalis 2006.) These foundational challenges strike me as tremendous for both research programs, and although some insightful work has already been done by way of addressing them, few philosophers are satisfied that the challenges have been fully met. At the same time, few philosophers take this as reason to suspend research within NERP, and the same attitude should attach, I contend, to PIRP.

⁵ Thus I find that PIRP has been hurt by its underdog status inasmuch as insufficient energy has been channeled toward its pursuit individually and regardless of the prospects of other research programs. Much more energy has been devoted to arguing that it *should* be pursued than to pursuing it. This is perhaps an understandable predicament for an ascendant research program in the initial phases, but there also comes a time where the program has to actually be pursued.

⁶ Furthermore, proponents of NERP and PIRP have tended to be quite confrontational, assuming that the two programs are in competition with one another and only one of them can be viable, but there are also ways of being more conciliatory. One way is to hold that a general theory of intentionality will have to incorporate elements from both programs. For example, according to Horgan and collaborators (2002, 2004), the kind of intentionality which can be the basis for all other intentionality must involve both a phenomenal character and a natural relation to external items. A different way to be conciliatory, which I prefer, is to hold that there are in fact two distinct properties, each with its legitimate claim on the term “intentionality,” such that one behaves roughly the way indicated in NERP and one roughly the way indicated in PIRP. Thus, in his response to Searle’s (1991) claim that intentionality depends essentially on consciousness, Davies (1995) argues that dependence on consciousness characterizes some kinds of aboutness but not others; which kinds of aboutness should be called “intentionality” may then be a merely verbal matter. My own inclination would be to call the NERP-ly property *representation* and reserve the term *intentionality* for the PIRP-ly property (the point of PIRP could then be put as the rejection of a “representational account of intentionality”). But if this labeling policy is deemed insufficiently neutral, I would be perfectly happy to use the labels “intentionality₁” and “intentionality₂” instead.

⁷ I say “implies” because although (PI) states that only conscious intentionality is underived, and this entails that all unconscious intentionality is derived. It would strictly *follow* that all unconscious intentionality derives *from* conscious intentionality only on the assumption that all derived intentionality derives *from* underived intentionality. This assumption is not tautologous, however, as far as I can tell. Thus, Bourget (forthcoming) seems committed to its denial.

⁸ The approach is naturalistic in that it denies that intentionality is a super-natural phenomenon and attempts to cast intentionality as part of the natural order, with its relatively familiar ingredients. It is not, however, comfortably classified as an instance of NERP as the latter was characterized in §1.

⁹ It is possible to avoid this infinite regress simply by advancing interpretivism as a non-reductive (or non-constitutive) account of intentionality (Child 1994). But the kind of interpretivism I want to offer as an account of *unconscious* intentionality is reductive, so I focus in this discussion on the reductive version. Anyway a non-reductive interpretivism is not really an *account* of intentionality – it is more like a *comment* on it.

¹⁰ For a paper-length development of the point made in this paragraph, see Kriegel 2010a.

¹¹ I mention something like this account as an option in passing in Kriegel 2003a and 2010a and expand on it in a more endorsing mode in Kriegel 2007.

¹² See Johnston (1989) for a more precise characterization. The key is this formula: “*x* is C iff *x* is such as to produce an *x*-directed response R in a group of subjects S under conditions K” (Johnston 1989: 145). When this biconditional holds *a priori*, then C is a response-dispositional property.

¹³ That is, for any unconscious intentional state, there is some ideal interpreter and some circumstances, such that the hierarchy of interpretations that state induces in that interpreter under those circumstances is capped at the first step. Thus we can conceive of an ideal interpreter all of whose mental states are conscious (but who does not have fewer mental states for that).

¹⁴ In some discussions of these matters, the term “interpretationism” is used to mean something like what I here call (following Goldman) “rationality theory.” Obviously this should not be confused with the

interpretivism I am defending here. The latter accounts for intentionality in terms of interpretation. The former offers an account of how interpretation works.

¹⁵ These are: theory; simulation; rationality; theory + simulation; simulation + rationality; theory + rationality; theory + simulation + rationality.

¹⁶ I expand on the mechanics of charity-based ascription of non-phenomenal intentional states in Kriegel 2010b.

¹⁷ This would take care of the derivation of the intentionality of paintings, photographs, etc., as well as the intentionality of traffic signs and other forms of symbols that combine linguistic and pictorial representations. If there are non-mental representations that are neither linguistic nor pictorial, nor based in stipulative conventions (which are also, and more straightforwardly, derivative from intentions), then the interpretivist would need to either show that they too can derive from Gricean intentions, or else provide an altogether new account of how they derive their intentional content from conscious interpretation.

¹⁸ This may have to be further tweaked, given that facts described as concerning unconscious intentionality may turn out to be reducible to facts describable in non-intentional (e.g., neurophysiological) terms. It could be insisted, of course, that such facts do not qualify as non-intentional facts, on the principle that a fact qualifies as intentional if there is an intentional description of it (and regardless of whether there is also, in addition, a non-intentional description of it). But alternatively, we could modify the characterization of the ideal interpreter by making explicit reference to the description under which the ideal interpreters knows the facts it does.

¹⁹ It may be thought that since the ideal interpreter is defined in terms of the *capacity* to draw inferences, and capacities can be inhibited, that there are some sub-ideal conditions for the ideal interpreter's execution of its interpretive function, namely, the conditions under which the relevant capacity is in fact inhibited. If so, we would need to specify in the response-dependent biconditional that the response is elicited in other conditions, conditions under which the interpreter's capacity to draw inferences is not inhibited.

²⁰ If this is not the case, and the ideal interpreter may sometimes ascribe to a phenomenal state a content different from the one the state carries in virtue of its phenomenal character, then the problem I discuss in this paragraph does not arise. It may be thought that another problem arises, namely, that there is tension between the interpretation-based ascription of content and the phenomenally based ascription. But this is not really a problem, since interpretivism put forward here is not an account of phenomenal intentionality, only of non-phenomenal intentionality. So any interpretation-based ascription is not constitutive of the content of phenomenal states.

²¹ For more on epistemic necessity, see Chalmers forthcoming.

²² The view could be interpreted as a kind of "response-dependent realism," akin to the kind of view McDowell (1985) has been pushing in meta-ethics, where moral properties are construed as secondary qualities but this is taken to be a form of realism. There are legitimate questions as to whether this is genuine realism and if so in what sense, but in any event the view is more realist in spirit than Dennett's.

²³ One reason to treat this particular test – consistency with intuition and existing scientific practice – as (quasi-)empirical is that it is a familiar thought in philosophical methodology that in the same sense in which empirical theories face the tribunal of experience, philosophical theories face the tribunal of intuition and of existing scientific practice. A philosophical theory does not make contact with any other sorts of data. The only data for philosophical theorizing are judgments produced either "intuitively" or on the basis of "our best theory of the world."

²⁴ In Kriegel forthcoming Ch.4, I argue that interpretivism is also more theoretically virtuous.

²⁵ It is not part of this claim that there can be no phenomenally unconscious occurrent thoughts. The claim is merely that, in addition to those, there are also phenomenally conscious occurrent thoughts.

²⁶ For defenses by more historical figures, see Husserl 1900 and Moore 1953.

²⁷ This view could be naturally augmented by the claim that there are also differences in the phenomenal intensity of attitudes: believing that p with great conviction is phenomenally different from believing that p with some doubt in one's heart.

²⁸ At the same time, it is important to note that (PI) can turn out to be true even if there is no such thing as cognitive phenomenology at all. Thus cognitive phenomenology is unnecessary for PIRP, though it would be greatly helpful, so to speak.

²⁹ For other arguments of this form, see Peacocke 1998, Horgan and Tienson 2002, Kriegel 2003b, Pitt 2004, and Chudnoff forthcoming.

³⁰ Similar arguments are presented in Goldman's (1993) and Mangan's (2001) discussion of tip-of-the-tongue phenomenology, Siewert's (1998, this volume) and Siewert's (this volume) and Chudnoff's (forthcoming) discussion of intellectual Gestalt shifts. The tip of the tongue experience is familiar to most of us and is claimed by Goldman and Mangan to involve a distinctively cognitive phenomenology. Siewert's delayed understanding concerns experiences in which we suddenly understand a piece of text we have been rereading and trying to understand for a while (imagine suddenly thinking of kites on your own when reading the above passage). Intellectual Gestalt shifts, meanwhile, concern experiences in which a phenomenon is construed one way but is suddenly reinterpreted another way; Siewert (this volume) illustrates this with an entertaining exchange he heard in Miami on an unbearably hot day, in which one woman said to the other "I am so hot" and the other responded "you don't have to brag about it."

³¹ Schematically, the argument proceeds as follows: we have a special, immediate access to our cognitive states (and their contents); only to conscious experiences (and contents) can we have this kind of special access; therefore, our cognitive states (and their contents) are conscious experiences (and experiential-intentional).

³² An earlier argument of a similar form is developed by Goldman (1993) and recently endorsed by Lycan (2008).

³³ I add a parenthetical mention of appearance because it is not meant to be built into this characterization of phenomenality that the explanatory gap is unbridgeable. Since some people hear the term "explanatory gap" as committing to the impossibility of bridging, and other as committing to the impossibility of a priori bridging, whereas I do not mean to characterize phenomenality in a way that commits to the existence of such an unbridgeable gap between it and physical property, I add the reference to an appearance.

³⁴ The renowned high-functioning autistic Temple Grandin reports (Grandin 1996) experiencing only four types of emotion: joy, sadness, fear, and anger. Subtler emotions are unfamiliar to her experientially, though she has managed to understand their nature in a purely intellectual, bloodless manner. This is an instance of congenital *restriction* of one's emotional phenomenology. We can readily envisage someone suffering from congenital *absence* of such phenomenology.

³⁵ So (a) there is no x which metaphysically depends upon itself, and (b) whenever some x metaphysically depends on y and y metaphysically depends on z , x metaphysically depends upon z .

³⁶ This is not to say that the intentional commonality underlies the phenomenal commonality rather than the other way around. In fact, it is not to comment at all on the order of explanation, or order of constitution, between the intentional and phenomenal commonalities.

³⁷ Here too, none of this prejudices either the order of explanation or the order of constitution between interpretive content and interpretive phenomenology. As I argue immediately in the text, for the proponent of PIRP the order goes from phenomenology to content: conscious interpretive acts have their interpretive content because of their interpretive phenomenology.

³⁸ It may be objected that this would work only if cognitive phenomenology is individuating in Pitt's (2004) sense, that is, if there is a different kind of cognitive phenomenology for every type of conscious cognitive state. My response is to simply accept the objector's claim and subscribe to the claim that cognitive phenomenology is individuating. As indicated in §2, this is not mandatory for holding that there is such a thing as cognitive phenomenology. But I accept the objector's claim that it *is* mandatory for holding that cognitive phenomenology is the basis of unconscious intentionality.

³⁹ Of course, the view does allow perceptual experiences with sensory phenomenology only. Perhaps a perception of a facial contortion is a case in point: an experience with a purely sensory perceptual phenomenology. But the point is that, in addition to such experiences, this conception allows also for more sophisticated perceptual experiences, ones whose phenomenology comprises a cognitive element.

⁴⁰ In a way, then, the difference between the view presented here and the historical view is the difference between making a claim about states and making a claim about properties of states. Also, the view presented here is uncommitted on the issue of factorizability, whereas the historical view is committed.

⁴¹ On some understandings of what it is for a state to be "cognitive," this may be a conceptual truth – namely, in case having propositional content is sufficient for qualifying as cognitive. However, even if we use another criterion for being cognitive, it strikes me as substantively true that perceptual propositional attitudes have a cognitive phenomenology (see, again, Klausen 2008).

⁴² Thus every phenomenally intentional state has a phenomenal character that involves phenomenal taking as a component, which is a cognitive component of its phenomenal character, and it is in virtue of this cognitive component that the state is intentional. This may raise the objection that certain animals and neonates may be robbed of intentionality. There are probably reasonable responses to this objection, but a full discussion would take us too far afield.

⁴³ Similarly, according to Georgalis, the essential subjectivity of conscious intentional state, in virtue of which they are such, is a "non-sensory" feature. Under certain assumptions (which Georgalis may or may not accept), this could be taken to mean that this essential subjectivity is a kind of cognitive phenomenology.

⁴⁴ For comments on a previous draft, I am indebted to David Chalmers, Nicholas Georgalis, Michelle Montague, David Pitt, two readers for OUP, and especially Tim Bayne. For very helpful conversations, I would like to thank George Graham, Terry Horgan, and Amie Thomasson. I also benefited from presenting the paper at a conference at the University of Bern, and would like to thank the audience there, in particular Michelle Montague, Philippe Keller, Sebastian Leuggers, Gianfranco Soldati, and Galen Strawson. The paper was partly written during a research fellowship at the University of Sydney, to which I am greatly indebted.