

Version read to the  
California Phenomenology Circle  
San Luis Obispo, June 1, 2007

## UNCONSCIOUS INTENTIONALITY

David Pitt  
California State University, Los Angeles

Some contemporary philosophers – most notably John Searle and Galen Strawson – have persisted in the Cartesian intuition that consciousness and intentionality are somehow essentially connected. Descartes himself held that the relation between consciousness and intentionality is especially intimate: there can be no unconscious mental states of *any* kind; and no creature incapable of consciousness is capable of mentality. Descartes's view is widely acknowledged to have been discredited by Freud, and by contemporary cognitive science. Freud argued that the best explanations for much of human behavior require that we suppose there are unconscious states, such as beliefs and desires, with intentional content. If the explanation of behavior requires us to advert to states with intentional content, then it requires us to advert to *unconscious* states with intentional content. Moreover, much of the explanatory success of cognitive science depends upon the supposition of “sub-personal” and otherwise unconscious mental representations. Contrary to what Descartes believed, a great deal – perhaps even most – of our intentional mental life goes on unnoticed by us.

Anyone who accepts the Freudian thesis, but who maintains that nonetheless there is a substantive sense in which intentionality is impossible without consciousness, is faced with the difficult task of reconciling these seemingly incompatible theses.

Both Searle and Strawson allow that intentional states can be unconscious. However, they both attempt to honor the Cartesian intuition by placing restrictions on the conditions under

which an unconscious state can be intentional. These restrictions create a strong tie between consciousness and intentionality, though (unlike Descartes) neither Searle nor Strawson requires of any particular state that it be conscious in order that it have intentional content. Searle (1991, 1992) claims that an unconscious state can have intentional content only if it could be conscious; and Strawson (1994, 2004) claims that an unconscious state can have intentional content only if its possessor could be conscious (that is, only if it is a state of a creature capable of consciousness). Hence, both claim that a capacity for states with intentional content depends upon the capacity for consciousness.

I do not think either Searle's or Strawson's position is stable, and in what follows I will try to establish this. I will defend the Cartesian intuition that consciousness and intentionality are essentially connected, however – though I will argue that the connection is even less direct than Searle and Strawson suppose. A mental state can be fully intentional even if it is not conscious, and even if it is a state of a creature that could not be conscious; yet the properties of a mental state in virtue of which it is intentional are properties that it cannot fail to have if it is conscious. They are of a kind that no conscious state of any kind could fail to have.

### **I. Searle's Thesis.**

Searle is fundamentally committed to the Cartesian intuition; yet, he accepts that there can be unconscious intentional states: “The explanatory power of the notion of the unconscious is so great that we cannot do without it” (1992: 151). Searle's position is an attempt at a compromise. He claims that “there are no *deep* unconscious intentional states” (id.: 162, emphasis added), where a *deep* unconscious state is one that cannot, in principle, be brought to consciousness (because it is just not the right *kind* of state), but that an unconscious state may be

intentional if it is, in principle,<sup>1</sup> potentially conscious. Thus, according to Searle, the property of an unconscious intentional state in virtue of which it is intentional is a property it would not have unless it were potentially conscious. But what property is that?

Searle's argument for an essential connection between intentionality and consciousness proceeds via his notion of "aspectual shape": "The link ... between intentionality and consciousness lies in the notion of an aspectual shape" (1991: 52). In my view, Searle nowhere gives a clear account of what aspectual shape is supposed to be,<sup>2</sup> but he makes the following claims about it. It is in virtue of having aspectual shape that a mental state has intentional content: "To be intentional, a state or process must be thinkable or experientiable; and to be thinkable or experientiable, it must have an aspectual shape" (id.). Aspectual shape is irreducibly *subjective*: "aspectual [shape] must exist from [the thinker's] point of view" (id.: 53). A state need not be conscious to be subjective: "there is something subjective about mental states even when they are unconscious" (id.: 56).

Aspectual shape properties, therefore, would appear to be the likely candidates for those properties in virtue of which an unconscious intentional state is intentional. For, the last two claims of the preceding paragraph together seem to entail that a state need not be conscious in order to have aspectual shape.

Yet, Searle *also* says that "The ontology of mental states, at the time they are unconscious, can only consist in the existence of purely neurophysiological phenomena" (id.: 53). Such "third person" facts cannot, however, determine an aspectual shape for a mental state: "There isn't any aspectual shape at the level of neurons and synapses" (id.: 59). So it looks like there is a flat-out contradiction here: an unconscious intentional state both has and does not have

aspectual shape. Searle comes closest to stating an explicit contradiction in his 1992 (161): “intentional states, conscious or unconscious, have aspectual shapes, and there is not aspectual shape at the level of neurons.”

Searle recognizes the problem: “now we seem to have a contradiction: the ontology of unconscious intentionality is entirely describable in third person, objective neurophysiological terms, but all the same the states are irreducibly subjective. How can this be?” (1991: 57). Of course, it cannot be. Searle attempts a compromise in the following way. He claims that an unconscious state may be intentional – may have aspectual shape – if it is *potentially* conscious: “When we characterize an unconscious intentional state in terms of its aspectual character [by which he must mean *when we characterize an unconscious state as intentional*], we are characterizing a present brain state in terms of its *causal capacity* to produce a conscious thought or experience” (id.: 58; emphasis added). Thus, Searle’s official position seems to be that the essential link between consciousness and intentionality lies in the fact that the property that makes an unconscious state intentional is its potential for consciousness – i.e., a causal power.

But this seems to imply an identification of the aspectual shape of an unconscious brain state with its power to cause aspectual shape, since it is the aspectual shape of a conscious state that makes it intentional. If Searle intends a distinction between unconscious aspectual shape and conscious aspectual shape, then he owes us an account of the difference. But, I shall now argue, there is no account of the difference that would meet all of the constraints Searle accepts. Searle is faced with two dilemmas, one embedded in the other. The main dilemma is this: Either unconscious aspectual shape is distinct from conscious aspectual shape, or it is not. If it is not, then Searle is forced into the incoherent position of identifying a property with a disposition to

manifest that property. If unconscious aspectual shape *is* distinct from conscious aspectual shape, then (this is the embedded dilemma) the difference is either sufficient to render it non-intentional, or it is not. If it *is* sufficient, the Freudian Thesis is no longer honored – unconscious states cannot have aspectual shape. But if it is *not* sufficient, Searle loses his thesis of an essential connection between consciousness and intentionality, for it would then be possible for unconscious states to be fully intentional, independently of any relations they might have to consciousness.<sup>3</sup>

In fact, I do not think it is entirely clear *what* Searle's position is.<sup>4</sup> It does seem to me, however, that none of the reasonably faithful interpretations of what he says presents him with a position he can accept.

## II. Strawson's Thesis.

In his book *Mental Reality*, Galen Strawson claims that “[t]here is a clear and fundamental sense in which *meaning*, and hence *intentionality*, *exists only in the conscious moment ...*” (209), and, hence, that “true or original or intrinsic intentionality is found only in the case of experiencing beings” (186-87). Nonetheless, he maintains, “[o]ne can acknowledge the sense in which meaning exists only in the conscious moment while endorsing a theory that attributes intentionality to some of the nonexperiential [i.e., nonconscious] states of experiencing beings” (209).<sup>5</sup> Only creatures capable of consciousness (“conscious creatures,” for short) can be in intentional states, but some unconscious states of such creatures may have intentional content. Thus, Strawson is honoring Freud's thesis, but limiting its application to conscious creatures.

Now, this position immediately forces the question, which Strawson recognizes, of how it is that the capacity for conscious experience could make such a difference. Why is it that

unconscious states of conscious creatures can have intentional content while those of non-conscious creatures cannot? Prima facie it would seem that unconscious states (physical, neural, dispositional, etc.) of the former could be intrinsically type-identical to those of the latter. So why could they not both have intentional content? More specifically, suppose that an unconscious state  $U$  of a conscious creature  $C^+$  has the intentional content *that p*. If it is possible for a non-conscious creature  $C^-$  to be in an unconscious state  $U'$  of a type identical to  $U$ , why shouldn't  $U'$  have the content *that p*?  $U$  and  $U'$  are intrinsically identical; why should their occurring in  $C^+$  and  $C^-$  make this difference?

I will consider two answers to this question, one explicitly addressed by Strawson, the other only suggested by some of what he says in chapter 7 of *Mental Reality*. I will argue that neither account provides adequate support for Strawson's claim that the capacity for consciousness is essentially connected to the capacity for intentionality. Given that Strawson does not offer or suggest any other account of the relationship between consciousness and intentionality, he has not succeeded in reconciling the Cartesian intuition with the Freudian thesis.

The first answer (1994: 205) is that only conscious creatures are capable of intentionality because (1) the property  $P$  that renders a creature capable of intentionality is a property only a creature could have, but (2)  $P$ 's being necessary for intentionality is *independent of* its being limited to conscious creatures; hence (3) though only conscious creatures can be in intentional states (because only they can instantiate  $P$ ), the reason  $P$  is necessary for intentionality is not the same as the reason it can only occur in conscious creatures. Perhaps more perspicuously:  $P$  has the second-order property  $P'$  of being instantiable only in conscious creatures, and the second-

order property  $P''$  of being necessary for intentionality, and  $P' \neq P''$ . Presumably,  $P$  has these distinct second-order properties in virtue of its structure: it is a compound of distinct first-order properties  $P_1$  and  $P_2$ , such that  $P_1$  can only be instantiated by conscious creatures, and  $P_2$  is necessary for intentionality. It is because of  $P_1$  that  $P$  can only be instantiated by conscious creatures; but it is because of  $P_2$  that  $P$  is necessary for intentionality. Hence, though the capacity for intentionality is distinct from the capacity for consciousness ( $P_1 \neq P_2$ ), any creature that has one will, necessarily, have the other, since they are “bundled” in a single property,  $P$ .

In these terms, the answer to the question posed above – Why should  $U'$  not have the same intentional content as the type-identical  $U$ ? – is that, contrary to what the question assumes,  $C'$  could not be in a state  $U'$  type-identical with  $U$ , because, given that  $U$  has intentional content, and that its having that content is essential to its type identity,  $C'$  would not be capable of being in it, since  $C'$  is not capable of consciousness. Since  $C'$  cannot instantiate  $P$  (because it cannot instantiate  $P_1$ ), it cannot instantiate  $P_2$  (because  $P_2$  comes bundled with  $P_1$ ), and, hence, none of its states can be intentional.

Now, Strawson rejects this proposal, and, I think, with good reason. Why, he asks, could there not be some property  $P^*$  which shares  $P_2$  with  $P$ , but not  $P_1$ ? If  $P_1$  and  $P_2$  are distinct properties, then, unless it is metaphysically necessary that only  $P$  has  $P_1$  as a constituent (which cannot be assumed without argument), then why could there not be a creature possessing  $P_2$  but not  $P_1$ ? Why indeed.<sup>6</sup> To insist otherwise – to insist that  $P_2$  could *only* be bundled with  $P_1$  – seems simply to be just another way to claim, without explanation, that the capacity for consciousness and the capacity for intentionality are essentially linked.<sup>7</sup>

On the second answer, which is not explicitly discussed by Strawson, it is the fact that the

unconscious state is in certain ways *related to* conscious states in the conscious creature that allows it to be a bearer of intentional content. Only conscious creatures are capable of intentionality because, though in the first instance “intentionality exists only in the conscious moment,” unconscious states that are part of *a network of states some of which are (or have been, or could be) conscious* can have intentional content. The idea is that, analogously to the way conceptual content is conferred by position in a network of concepts on functionalist, inferential-role and conceptual-role theories of content, the possibility of intentionality is conferred upon an unconscious state in virtue of its relations to other states, some of which are, at some time, conscious. No member of a network of states none of which was ever conscious can have intentional content, though unconscious states in a network some of whose members are at some time conscious can have intentional content.

It is not clear, however, *why* this would be. Why should inclusion in such a network confer a capacity for intentional content? It is not plausible that this is simply a brute fact. But if there is some *intrinsic* property, distinct from consciousness, that may be conferred upon an unconscious state in virtue of its relation to a conscious state,<sup>8</sup> and in virtue of which the unconscious state can have intentional content, then why should such a property not be conferrable, at least in principle, by some other, *non-conscious* source? Again, the problem is that to separate consciousness-conferring properties and intentionality-conferring properties is to open up the possibility that, since they are distinct, they may be instantiated separately in organisms generally.

Neither of these accounts is satisfactory. So why does Strawson think that the capacity for consciousness confers the capacity for intentionality? He says: “Having allowed that

nonexperiential states can be intentional states, one may go on to consider whether entirely experienceless entities can be said to be in intentional states. I think not, but it is not as if there is a simple right answer. It depends on how one interprets the word ‘intentionality’ ” (186-87). *True* intentionality, Strawson says, is only found in experiencing beings; though one may find its ersatz – a property constituted by microbehavioral dispositions to respond to environmental stimuli, what Strawson calls ‘behavioral intentionality’ – in the inveterately senseless. But Strawson is not interested in the ersatz; he is after the Real McCoy.

So what is Strawson’s view? As far as I can tell, he maintains a stoic silence. Having raised the problem, he leaves it as such. In doing so he has left out his most important argument, however, and has not presented us with a position we have reason to accept. I share with Strawson the intuition that consciousness and intentionality are somehow essentially connected; and I agree with him that the intuition is a very strong one, and one worth exploring. But intuitions are notoriously idiosyncratic and subject to theoretical infection, and they too easily bring us to loggerheads. One ought to have cogent explanations of what one finds intuitively compelling.

Neither Strawson’s nor Searle’s account succeeds in reconciling Descartes and Freud. In what follows I will sketch and defend an account that does. On this view, intentional states need be neither conscious nor states of a conscious creature, yet there remains a deep connection between intentionality and consciousness: though intentionality does not require consciousness, it does require properties of a kind which *any* conscious state must have. Intentionality and consciousness have a common necessary condition, namely, *phenomenality*.

### **III. Intentional Phenomenology.**

I have argued elsewhere (Pitt 2004) that each conscious thought has a phenomenology of a kind that (i) is different from that of all other types of phenomenal states (bodily sensations, perceptual sensations, images, moods, etc.), (ii) is different from that of any other conscious thought, and (iii) constitutes its representational content.<sup>9</sup> Conscious thoughts on this view have phenomenologies that are (i) *proprietary*, (ii) *distinctive*, and (iii) *individuating*: the phenomenology of the conscious thought that *p* distinguishes it from all other conscious mental particulars – including all other conscious thoughts – and *makes it* the thought that *p*.<sup>10</sup>

Here, I wish to argue, further, that phenomenality is conceptually, metaphysically and nomologically distinct from consciousness: the concept of phenomenality is not the concept of consciousness; phenomenality is not the same property as consciousness; and it is nomologically possible for phenomenality to be instantiated without consciousness being instantiated. Yet, though phenomenality and consciousness are not the same, it *is* metaphysically (hence, nomologically) impossible for consciousness to be instantiated without phenomenality being instantiated. It is not possible to be conscious without being conscious *in some way*, and the *ways* of being conscious are just phenomenal properties. Hence, though phenomenality does not require consciousness, consciousness does require phenomenality. Since (on my view) intentionality also requires phenomenality, there is an essential, though indirect, connection between consciousness and intentionality: they have a common necessary condition. The existence of unconscious thoughts is thus compatible with the Cartesian intuition, though not in the way that either Searle or Strawson claims.<sup>11</sup>

I do not think it can be the case the phenomenality just *is* consciousness. There are so many *kinds* of phenomenology, so many *ways* of being conscious – many of which are *radically*

different from each other – all of which are (or can be) conscious *in exactly the same sense*.<sup>12</sup> To say that an experience of pain is conscious and that an experience of bright yellow is conscious is to say the same thing about them: they are such that a subject is aware of them, and of what they are like. Qua *conscious* there is no difference between them.<sup>13</sup> Even if one insists that the forms of phenomenology are *just* forms of consciousness, it is still the case that consciousness is a different *property* than phenomenality. And even if a state cannot possibly be phenomenal unless it is conscious, it is still the case that consciousness and phenomenality are different properties.

Perhaps the most frequently offered evidence for the nomological independence of consciousness and qualia is the *unconscious qualia* interpretation of blindsight cases.<sup>14</sup> I think this interpretation is quite compelling, though I also think it has not been sufficiently developed or defended.

In Weiskrantz's famous experiments, subjects with a type of brain damage that resulted in their having blind spots ("scotoma") in their field of vision – areas within the visual field in which there is a complete lack of conscious visual experience – nonetheless responded surprisingly accurately when asked to guess at stimuli presented in the blind field. For example, if asked whether or not a light had been flashed in the blind area, patients will guess correctly at rates far above chance. Even more striking are cases in which patients are able to guess at the completion of a figure, half of which is consciously experienced and half of which is within the blind field, and hence not consciously experienced.

Some philosophers have wanted to claim that this shows there can be unconscious sensations – i.e., unconscious experiences with qualitative character. There is, however, an

obvious alternative explanation of blindsight patients' capacities which does not posit the existence of unconscious qualia. Namely, that non-phenomenal *information* about the properties of the stimulus can in these patients still get through to a central processor (or whatever). Thus, though there might be some problem with a blindsighter's visual cortex, his rods, cones, retinas, optic nerves, and all manner of other apparatus "up to" (so to speak) the visual cortex, are in normal working order. And whatever connections there may be between lower-level differential activity in one's visual system and what gets tokened in one's "belief box" are intact in these patients, though the connections to whatever mechanisms it is that generate conscious, phenomenal experience are disrupted.

Is there anything that can be said in response to this? I think there is. One needs a clear idea of just what sort of *work* phenomenology does – that is, some idea of what someone lacking a certain kind of phenomenology could not do that someone possessing it could do. For if we had a clear idea of that, then we might be able to say something about what blindsighters can do that those completely lacking phenomenology cannot do, and side-step the whole question of consciousness.

Consider those who lack, not consciousness, but phenomenology of a certain kind – for example, people with *achromatopsia*. Achromats lack, to varying degrees, color vision; *complete* achromats see only in black and white and shades of grey. Their visual experiences of things are thus, phenomenologically, very different from those of the normally sighted. They may be incapable of, or have great difficulty in, identifying or distinguishing things on the basis of their colors. Red things and black things, for example, may look the same to them; whereas to those who are not achromats the experience of seeing a red thing is quite different,

phenomenologically, from the experience of seeing a black thing. What is the significance of the difference? Just this: normally sighted individuals can *distinguish* red and black things on the basis of their color, and achromats cannot. That is, it seems that it is the *phenomenology* that makes the difference here, for, we may assume, achromats are otherwise phenomenologically and cognitively just like the normally sighted – including being equally *conscious*.

Now, if blindsighters could make distinctions on the basis of color that the achromat could not, I think we would have a stronger (though not, of course, conclusive) reason for thinking that the phenomenon of blindsight – or at least some cases of it – do show that there can be unconscious phenomenology. But it seems that they can. Dennett (1991), for example, cites an experiment, reported in Stoerig and Cowey 1990 that, he says, provides evidence that it is possible for blindsighters to discriminate colors in their blind field. But if this is an ability that depends on there being for the subject phenomenological differences between the stimuli, it looks like blindsighters nonetheless experience phenomenology, though unconsciously.

Compare the conscious and fully sighted to the conscious though achromatopic to blindsighters to those who are completely unconscious. Blindsighters can do things the completely unconscious cannot; conscious achromats can (one assumes) can do things the blindsighted cannot; and, of course, the fully sighted and fully conscious can do still more. But notice that this progression seems to suggest a progressive building up of capacity – a progression that suggests that the capacity for consciousness and the capacity for phenomenology are distinct. The achromats compare to the “chromats” with respect to phenomenological capacity; the blindsighted compare to the fully sighted chromats with respect to capacity for consciousness, but not with respect to capacity for phenomenology; and the fully unconscious

compare, rather unfavorably, in all respects to all of the others.

But why not offer the same explanation for the blindsighters' ability as before? Why not say, that is, that what gets through is nonphenomenological information, and go on to claim that it is just *that* that is missing from the achromats?

I think there are reasons – offered by Raffman (1995) – to think that this could not be true in the general case. And if it could not be true in the general case, then perhaps the very cases that Raffman discusses could be implicated in blindsight cases. Raffman argues, convincingly in my view, that we are able to make phenomenological distinctions that outstrip our capacity for making conceptual distinctions, and that, therefore, there is a distinctively phenomenal sort of information that we are capable of possessing.<sup>15</sup> If there is such information, and it is possible for phenomenological states to be unconscious, then it is at least possible that there could be cases of blindsight discrimination of perceptual stimuli that are made on the basis of phenomenology alone.

In general, *if phenomenality does anything at all*, it is to enable us to make certain distinctions perceptually. Perhaps something else *could* do what phenomenality does; but nothing else does it *for us*. It is easy to imagine creatures that could make all the perceptual discriminations we do, who were as adept at navigating their environments as we are, and who are as attuned to their inner lives as we are, but who have no phenomenal states at all. But if you deprive *us* of phenomenality we suffer drastic reductions, if not extinctions, in these areas. It is, therefore, not unreasonable (though certainly not *necessary*) to think that it is *phenomenality* that enables us to do these things, and not the non-phenomenal, unconscious neural processing that remains when phenomenality (and, hence, consciousness) is removed. Those nonphenomenal

creatures who are as competent as we are would have to have much more sophisticated nonphenomenal mechanisms than we do. It may be that evolving phenomenality is easier (or maybe just easier on Earth), or that it is an historical accident that these various functions are subserved by phenomenality in us; but I do not think it can reasonably be denied that phenomenality does what it does for us. And if we can still do some of what conscious phenomenology allows us to do when consciousness is removed, then it is not unreasonable to suppose that it is still phenomenology that explains the residual abilities.

I think it is not unreasonable, therefore, to conclude that the phenomenon of blindsight (and perhaps other phenomena) does provide evidence for the claim that phenomenology in general need not be conscious.<sup>16</sup> If this is so, then there is a basis for claiming that *cognitive* phenomenology need not be conscious, and, hence, that the connection between consciousness and intentionality is their shared dependence upon phenomenality.

#### **IV. Consciousness, Subjectivity and the Self.**

Though I think the unconscious-phenomenology interpretation of blindsight cases is a reasonable one, I would nonetheless like to briefly sketch a version of Descartes's view, on which there can be no unconscious intentional states (because there can be no unconscious phenomenology), but on which unconsciousness need not be an *intrinsic* property of a mental state. On this view, a state that is conscious *in itself* might nonetheless be unconscious *for its possessor*. Complex issues regarding the nature of the self, of subjectivity, and their relations to consciousness and phenomenality are lurking at the edges of this view; but there is not space here to address them. (I plan to do so elsewhere.) All I hope to do here is to make the idea of a state conscious in itself but not for its possessor a bit more plausible than it might at first seem.

Could I be in a conscious state that I am not conscious of? That is, could a state conscious in itself, but of which I am not directly aware, be *mine*? Surely there are states conscious in themselves that I am not (directly) aware of – viz., everybody *else*'s conscious states. But could such a state be *mine*? Is a state of mine conscious if and only if it is conscious *for me*? To ask this is not to ask if a state is conscious if and only if it is conscious for *someone*. Maybe this is true (maybe consciousness presupposes a *self*). The question I am asking is whether when a conscious state is *mine* that someone has to be *me*. I think the answer is no. I think it is possible for me to be in a conscious state of which I am not directly aware – i.e., for there to be a state of mine which is *conscious-in-itself* but not *conscious-for-me*<sup>17</sup> – and, further, that this provides a way of understanding unconscious intentionality consistent with Descartes's view of the relationship between consciousness and intentionality: a state is intentional only if it is conscious, but an individual need not be directly aware of all of his conscious intentional states. But how could this be?

Think of your own consciousness – your own direct awareness – in relation to the conscious states of another individual – call her Penelope. You are not conscious (directly aware) of Penelope's conscious thoughts: they are not part of your conscious experience. Why not? Presumably, at least in part, because you are physically distinct. Your brain and central nervous system produce your conscious experience, her brain and central nervous system produce hers, and the two are distinct, self-contained systems. But perhaps they could be interconnected in such a way that you and Penelope would become directly aware of each other's conscious thoughts, as they occur. Then we would have (maybe<sup>18</sup>) a case of your being *directly* aware of someone else's conscious states. (Likewise, of course, for Penelope.) Your relation to

Penelope's conscious thoughts would be just like your relation to your own conscious thoughts, insofar as you are directly aware of both. Hence, though Penelope's conscious thoughts originate in her brain, and yours originate in yours, and in *that* sense are the thoughts of distinct individuals, there is another sense in which Penelope's thoughts would also be yours (and yours hers). For a conscious state to be *mine* in this sense is for it to be an object of my direct awareness.

Now imagine that you and Penelope have been connected in such a way that her conscious thoughts can figure in the etiology of your behavior. Sometimes you do things because of what you consciously believe and desire; sometimes you do things because of what she consciously believes and desires. Thus, the explanation of your behavior would occasionally have to refer to Penelope's conscious states. It seems to me that this is yet another sense in which Penelope's conscious thoughts could also be yours. When we attribute intentional states *of which they are not directly aware* to agents, it is in order to explain their behavior. The agent may deny that the state is *his*, because he is not conscious of it (not directly aware of it); yet we might insist that it nonetheless *is* his, since the best explanation of his behavior requires attributing it to him.<sup>19</sup>

Hence, though Penelope's thoughts are hers in the sense that they originate in her brain and not yours, there are two ways in which we might be tempted to say that they are also yours: you may be directly aware of them, and they may figure in the causation of your behavior.

Now suppose that we alter your connection to Penelope in such a way that you are no longer directly aware of her conscious thoughts, though they continue to affect your behavior. Then we would have a situation in which Penelope's conscious thoughts would be yours in the

second sense, but not yours in the first sense: you are no longer directly aware of them, but they continue to figure in the causation of your behavior. Of course, Penelope's conscious thoughts are still hers in the sense that they originate in her brain, not yours. But a bit more surgery could change that.

Suppose we had, unbeknownst to you, transplanted Penelope's brain, hooked up to yours in the same way as in the previous paragraph, into your skull. You are now a *double-brain* patient. You have (we may suppose) two distinct centers of consciousness in your head, both of which have an influence on your behavior, but only one of which *you* are directly aware of. Given that Penelope's brain is now housed in your skull, I think there is reason to say that you now have two brains.<sup>20</sup> You would be in significant ways like a *split-brain* patient. Such individuals also appear to host two distinct centers of consciousness in their skulls; yet we say these consciousnesses are both *theirs*. Moreover, we also say of them that their behaviors are theirs, though there is a detectable split in their etiologies: some behaviors are controlled by the right hemisphere and some are controlled by the left. If these are indeed two distinct centers of consciousness and control, the fact that they coexist in the same skull and exercise control over the same body nonetheless prompts us to attribute the conscious thoughts that figure in the etiology of the individual's behavior to the *same* individual. In these ways, your brain and Penelope's brain have become just like the two hemispheres of a split-brain patient – though instead of resulting from a fission, they result from a fusion. Hence, if Penelope's brain is now your brain, then her conscious thoughts would also be your thoughts in the sense that they arise from your brain.

Of course the split-brain patient's separate consciousnesses were once one, and his

hemispheres have a common history, whereas yours and Penelope's consciousnesses need never have been shared (we could have skipped the first stage of the imagined procedure), and your brains have separate histories. And this might be taken to be (though I doubt that it is) sufficient reason to resist accepting that Penelope's brain is now yours (or a part of yours). But the point I wish to make is that there is at least *one* substantive sense in which Penelope's conscious thoughts, though you are not directly aware of them, are *yours*: they have a direct effect on your behavior. We would have to refer to conscious states of the second brain in order to explain some of what you do. And these states, conscious in themselves, would not be conscious *for you* – you would not be directly aware of them.

The moral of this thought experiment is this. It does not seem impossible that there could be, in my head, thoughts which are conscious in themselves, which have direct effects on my behavior, but of which I am not directly aware – thoughts conscious in themselves, but not conscious for me, and yet still mine: thoughts which are simultaneously conscious and unconscious. They are conscious in the sense that they have phenomenal character (where this is thought of as *entailing* consciousness); but they are unconscious in the sense that I am not directly aware of them. If Penelope's (former?) brain could produce such thoughts, and a split-brainer's hemispheres could too (the conscious thoughts of each are, to the other, unconscious, though all are thoughts of the same individual), then why couldn't some part(s) of a single, intact brain do it? Perhaps we all have a Penelope – or a multitude of (perhaps permanent, perhaps momentary) Penelopes – in our heads, all of them capable of exerting some influence over our behavior, though none experienced as part of *us*.<sup>21</sup> Perhaps the limits of my conscious self – what I can be directly aware of – are not the limits of my consciousness, and what we call

unconscious thoughts are in fact conscious thoughts we are not conscious of. We are not conscious of them, not because *they* are unconscious, but because *we* are not conscious of them. For a thought to be unconscious is on this view for it to be unconscious *for us*, not *in itself* – no thought is unconscious in itself. In this way, Descartes and Freud could be reconciled: there is no thought without consciousness; but not all conscious thoughts exerting a direct influence on behavior are available to the conscious self.

Or maybe there's just unconscious phenomenology.

### V. Conclusion.

According to the view defended in (the first four sections of) this paper, consciousness and intentionality are essentially, though indirectly, linked. A state cannot be intentional without being phenomenal, since intentional content is (or is determined by) cognitive phenomenology. And a state cannot be conscious without being phenomenal, since phenomenal properties are just ways of being conscious. But an intentional state need not be conscious, since phenomenology need not be conscious. Unconscious phenomenal states are *possible* contents of consciousness (Searle), because phenomenal properties are possible ways of being conscious. And though it need not be true that a creature must be capable of consciousness in order to be capable of intentionality (Strawson), because phenomenal properties need not be conscious, it is not difficult to see why one would think otherwise, if one believed that “understanding experience” (cognitive phenomenology) has something to do with intentional content.

This view makes available an interesting alternative to the three major types of account of the nature of mentality that take experience seriously – the *Cartesian*, the *Brentanian* and the *Humean*. On a Cartesian view, like Searle's or Strawson's, *consciousness* is the defining feature

of the mental: creatures not capable of consciousness are not capable of mentality. But not all mental states are conscious, and (arguably) not all creatures incapable of consciousness are incapable of mentality. So the Cartesian view does not adequately characterize what is distinctive of mentality. On a Brentanian view (such as Dretske's, Lycan's or Tye's reductive representationalism<sup>22</sup>), *intentionality* is seen as the defining property of the mental: a state is a mental if and only if it is *about* something, and a creature not capable of such states is not capable of mentality. But (arguably), not all mental states are intentional (e.g. simple sensations), and it seems possible for there to be individuals (e.g. Davidsonian "swamp" creatures, lifelong brains-in-vats, Ganzfelders<sup>23</sup>) capable of experience but not of intentionality. So the Brentanian view does not provide an adequate characterization of the nature of mentality either. On a Humean view, *sensationality* is the defining feature of mentality: the contents of the mind can be divided into sensations (sensory and introspective experiences) and copies of sensations (images), and a creature incapable of inner or outer experience would not be capable of mentality. But (arguably) not all mental states are sensational (e.g., intentional states are not), and it seems possible for there to be *purely* cognitive beings (God?). So the Humean view also fails to capture what is distinctive of the mental.

On the assumption that there is a distinctive sort of cognitive-intentional phenomenology, together with the possibility that phenomenology in general need not be conscious, and the assumption that sensational and cognitive states are all the mental states there are, there is another way to characterize what distinguishes the mental from everything else. What thoughts, conscious and unconscious, and sensations, conscious and unconscious (and any other states, e.g., emotions or perceptions, that might be constructable out of them) all have in common is that

the are *phenomenal*. I think Strawson is right when he says (1994: xi) that “the only distinctively mental phenomena are the phenomena of conscious experience,” but that the implication that the phenomena of consciousness are essentially *conscious* is incorrect. What makes the phenomena of conscious experience – sensations, perceptions, thoughts, emotions, moods, etc. – mental is their having phenomenal properties, not their being conscious. The phenomena of consciousness are, qua mental, essentially phenomenal. *Phenomenality* is the mark of the mental.<sup>24</sup>

## NOTES

1. There may be reasons why, as a matter of fact, it may never become conscious (repression, etc.), but it is the right *kind of* state to be conscious.
2. In his 1992, Searle explains that “Whenever we perceive anything or think about anything, we always do so under some aspects and not others”; we “have a conscious experience of the object from a certain point of view and with certain features” (156-57). But this is not especially illuminating. In my 2004 (6), I suggest that *cognitive phenomenology* – a distinctive, content-determining “what it’s like” of conscious thoughts – may play the role Searle assigns aspectual shape.
3. This is, essentially, Armstrong’s objection to Searle (see Armstrong 1991).
4. Lepore and Fodor (1994) are also at a loss to understand what Searle thinks the connection between intentionality and consciousness is. It does seem to me, however, that their trouble is due more to a lack of sympathy with the Cartesian intuition than with any difficulty in interpreting what Searle actually says.
5. Strawson presents an updated statement of his position in Strawson 2004.
6. This is, again, essentially Armstrong’s objection to Searle.
7. A similar objection would apply to an account on which the instantiation of  $P_i$  (and, hence,  $P$ ) is not just *limited to* consciousness-capable creatures, but in fact *bestows* the capacity for consciousness on the creatures that instantiate it, as well as to an account on which the relevant properties were properties of *states* rather than *creatures*.
8. More specifically (and less plausibly), to a state which may *have been* conscious or which *might yet be* conscious. Can a creature which has tokened, say,  $10^6$  distinct mental states, only one of which was conscious, 35 years ago, have a state with intentional content?
9. The *representational* content of a thought is those of its intrinsic features that are relevant to its representing the proposition it does. The *propositional* content of a thought is the proposition (the mind- and language-independent abstract object) that it represents. Thoughts with the same representational content might represent different propositions; and the same proposition might be represented by different representational contents. In more recent work (“Intentional Psychologism”), I have explored a view on which intentional phenomenal types *are* the contents of thoughts.
10. The argument for this view is, in brief, that we can introspectively and non-inferentially distinguish our occurrent conscious thoughts from other occurrent conscious mental particulars, distinguish our occurrent conscious thoughts from each other, and identify our occurrent conscious thoughts as the thoughts they are (i.e., as having the intentional contents they do), and that we would not be able to do these things if conscious thoughts did not have phenomenologies

that are proprietary, distinctive and individuating.

11. Nonetheless, Searle's intuition that an unconscious state must be of a *kind* that is potentially conscious in order to have intentional content is honored by this view in its claim that conscious states are necessarily phenomenological. For if phenomenality is a condition of consciousness, then if a state has phenomenality it is a potential content of consciousness: it is the *kind of* state that could be conscious.

12. By 'conscious' I mean, to use Block's (1995) terminology, *phenomenally* conscious – i.e., conscious in the way a state is conscious when a subject is aware of being in it, and of what it is like to be in it.

13. Perhaps consciousness can come in *degrees*, so that one state (the pain) might be *more* or *less* conscious than another (the bright yellow) (or vary in consciousness itself over time). But this need not, per se, constitute a difference in the intrinsic *pain* and *yellow* phenomenal characters. In which case degrees of consciousness could be safely set aside here. On the other hand, if there are varying degrees of consciousness, and *how* conscious a state is affects its (total) phenomenal character (or, even if there are not, and it does not), perhaps consciousness *itself* is a species of phenomenology, alongside all the others. (Perhaps it is something like experiential *illumination* – in contrast to, e.g., *pain*, experiential *yellow*, experiential *loudness*, etc.). It would then perhaps be easier to see how a state could be phenomenal without being conscious: it is the same sort of thing as a state being phenomenal without being, say, *painful*. (Though then it might not be easy to see why a state could not be conscious without being phenomenal in some *other* way – i.e., why one cannot be *simply* conscious.) I cannot pursue these intriguing suggestions here.

14. Other phenomena that are sometimes taken to show that there can be unconscious qualitative experience are *subliminal perception* and *unexperienced pain* (e.g., the headache that persists through a time during which it is not felt, or the pain that wakes one from a deep (presumably dreamless) sleep). (Cf. Rosenthal 1991.) The fact that these hypotheses are not conceptually incoherent shows that the concept of consciousness and the concept of phenomenality (or experience) are not the same concept.

15. Peacocke (1983) makes similar claims, though I think he has since revised his view somewhat.

16. There also appears to be neurological evidence that consciousness and phenomenality are subserved by distinct structures (“dorsal” and “ventral streams”) in the brain. See Block 2001 and references therein.

17. The distinction is similar to Block's distinction between access consciousness and phenomenal consciousness (Block 1997); though I do not suppose that Block's distinction exhausts what one might mean by consciousness. Consciousness might be a primitive property, identical with neither phenomenality nor accessibility.

18. One might want to insist that if you are directly aware of Penelope's conscious thoughts then they are also your thoughts – that you and she have somehow *merged* or *melded* into one consciousness. I don't think so; but it won't make a difference to what I ultimately want to say.

19. I do not think the second and third senses of mental state ownership should be conflated. For it seems possible for there to be states which arise in my brain (they are mine in the second sense) but which have no influence on my behavior (they are not mine in the third sense) – either because they figure in the causation of no behavior at all (I'm completely paralyzed), or because they figure only in the causation of someone else's behavior (through some other kind of sci-fi hookup).

20. If we had transplanted Penelope's heart into your chest after removing yours, would we not say that *you* had a new heart – that the heart in your chest was *yours*? And what if we had left yours in?

21. Or perhaps consciousness does not require a *self* at all. I have been assuming that it does; but perhaps there can be subjectless conscious states. Then a view of the sort sketched here would not have to populate our heads with any *others*. I am inclined to think that consciousness and subjectivity are not intrinsically connected, though I cannot argue for it here.

22. See Dretske 1995, Lycan 1996, Tye 2000.

23. A made-up term for creatures capable only of homogenous visual experience, one color at a time.

24. Thanks to Heather Gert, Charles Siewert, Lynn Stephens, Galen Strawson and audiences at the 200X Tucson Consciousness Conference and California State University, Los Angeles for helpful discussion and criticism.