

The Representational Base of Consciousness

Andrew Brook

Institute of Cognitive Science, Carleton University, Ottawa, ON K1S 5B6, Canada

Paul Raymont

Department of Philosophy, Trent University, Peterborough, ON K9J 7B8, Canada

1. Abstract/Introduction

Current views of consciousness can be divided by whether the theorist accepts or rejects cognitivism about consciousness. Cognitivism as we understand it is the view that consciousness is just a form of representation or an information-processing property of a system that has representations or perhaps both. Anti-cognitivists deny this, appealing to thought experiments about inverted spectra, zombies and the like to argue that consciousness could change while nothing cognitive or representational changes. Nearly everyone agrees, however, that consciousness has a *representational base*. Whether consciousness *simply is* representational or cognitive, it at least *requires* representation (and cognition). In an ecumenical spirit, we will focus on this point of agreement and sketch a theory of what this representational base might be. We hope that the result will be a framework useful for investigating consciousness empirically.

There have been few developed attempts to say what this representational base might be like. Perhaps the best-developed are what is sometimes called the first-order representation approach (FOR) (Dretske 1995 and many others), in which conscious states are viewed not as objects of representation but as things one knows about by inference from the content represented in them, and the higher-order representation model (HOR) (Rosenthal 1997 and many others), in which the representational base of consciousness is a thought directed at one's own psychological state(s). Both approaches face serious difficulties. We propose an alternative, based on a somewhat radical version of the idea that representations are *self-presenting*.

2. FOR and HOR

FOR models take many forms, but the key idea in the variant that we will consider is this. When we represent something consciously, we are directly conscious only of what our state represents. We are directly conscious *via* a conscious state but not *of* it. Consciousness *of* it is not direct; it's an inference, an especially secure inference, from the fact that we are conscious of what is represented.¹ All we know about our representing is what we can infer from the ways in which the represented items appear. Dretske calls this inferential view of

¹ As Dretske (1995, p. 56) puts it, "You cannot represent something as F without, necessarily, occupying a state that carries the information that it is F (not G or H) that you are representing something as." (See also pp. 61-2.)

consciousness of representations displaced perception.

If true, FOR would have obvious attractions. For example, we could stop fussing about qualia and higher-order representations, since neither would play any role in consciousness. Unfortunately, FOR faces some problems.

First, consider itches, pains and other bodily sensations. On FOR, how it feels to have a pain is exhausted by what it represents – some bodily damage or the like (Tye 1995). But pains, itches and so on do not seem to represent anything (or anything beyond themselves, a point to which we will return), and even if they do represent something else, pains hurt and that is more than a matter of representing some bodily state.

Second, in general, when I am conscious of a representation by having it, this consciousness seems, contrary to FOR, to be as direct and *non-inferential* as any consciousness could be. The idea that we become conscious of pains and itches only via an inference seems just as implausible, and in exactly the same way, as the Dennettian idea that pains are a result of interpretation.²

Third, when I am conscious of something, I am conscious not only of *what* is represented but also of *how* it is represented. Some aspects of how an object is represented (visually, aurally, etc.) seem to be aspects of the representation itself. We readily discern such differences – between, for example, seeing a corner and feeling it, not to mention imagining it, remembering it, and so on. How could FOR account for such differences? It has to say that we come by this knowledge by way of inference – and that seems simply wrong. In general, FOR has trouble dealing with situations in which there is one content but two or more ways of representing it.

Finally, what about altered states of consciousness (via meditation, drugs, etc.)? It seems even less plausible than in the case of pains and so on to say that these states are about anything beyond themselves. If so, what is there to infer their nature *from*? Yet we are vividly conscious of them. Even if it turns out that altered states of consciousness or the others must be about *something* beyond themselves, what it is like to have those states can change without how other things appear to us in them changing.

The upshot is that FOR is almost certainly false. We are directly conscious not only *via* conscious representations but also *of* them.

HOR-theorists might agree with parts of our critique of FOR, but their alternative faces even worse problems. The key idea behind HOR, again, is that what makes me conscious of, say, a pain is something *distinct* from the pain, namely, a higher-order representation of it.

The worst problems facing HOR are due to the fact that it separates the representing state that confers consciousness (a higher-order state) from the represented state (e.g., a pain) on which consciousness is conferred. For the most part, representations can exist in the absence of what they represent. Suppose that the represented state isn't real – I *imagine* myself perceiving something, for example. For HOR, thinking about a real perception makes me conscious of the perception. Clearly, imagining a perception can also make me conscious of something (the imagined perception) – without, necessarily, me being conscious of the act of

² Dretske recognizes that he has a problem about pains, etc.: “this is a topic that I have neither the time nor (I admit) the resources to effectively pursue.” (1995, p. 103). He says much the same about depression (p. xv).

imagination itself. But the only real state in play here is the act of imagination (the imagined state being *merely* imaginary). If so, the act of imagination has to be the conscious state (even though, on HOR, I would not be conscious *of* it). With no lower state, nothing could be higher, so here we have consciousness with no higher-order state.³

When the represented item does not exist, things will seem to the subject just as they would if the represented item (e.g., a pain) had been real.⁴ This is because how things seem to the subject is determined by the representing state, by how it represents things as being. Thus it may seem to the subject for all the world as if she really were in a conscious state, such as a pain, even though there is no such state – indeed, by HOR she could even be in no conscious state at all at the time. This, we think, is a *reductio* of the HOR approach. If things seem to me just as they would if I were really feeling a pain, I must at least be in a conscious state.⁵

An interesting implication of these arguments is this. Contrary to FOR, the subject is directly conscious of her pain, a conscious state; contrary to HOR, she is made thus conscious by that very same state and not via another representation. If so, it follows that such conscious states are self-presenting.⁶ The following objection will suggest itself: ‘All you have shown is that representations must present their own *contents*, their own *object* – the imagined perception for instance. You have not shown that they must present themselves.’ Answering this objection will reveal just how dramatically deficient HOR is. Here is our response.

Take a conscious state C. There are only two options. The representation of C is either a distinct, higher-order mental state, or it is C itself (in which case C is self-representing). The arguments we mounted against HOR entail that the first disjunct is false, i.e., shows that the representation of C is not a distinct mental state.

First argument: If one became conscious of C by a distinct representation R of C, then how C seems at t will be fixed by how C is represented at t by R. If so, as long as R is present, things will seem just as they would if one really were in C, even if C were absent. For example, suppose C is a pain; then it could seem for all the world as if one were in pain – without any pain! Since this is not possible, if HOR entails that it is possible, and HOR does, this is a *reductio* of the first disjunct.

Second argument: If one became conscious of C by a distinct representation R of C, as long as R is present, things will seem to me just as they would if I really were in a conscious state C *even if I'm in no conscious state at all at the time*. Since this too is not possible, if HOR entails that it is possible, and HOR does, this too is a *reductio* of the first disjunct.

From which it follows that the representation via which one becomes conscious not just of the contents of C but of C itself has to be none other than C. That is to say, C is and has to be self-presenting. One might not realize that it is – one might not think of it *as* something that is self-presenting – but in fact it *is* self-presenting. Equally, one might *also* have higher-order

³ The conscious state will of course have embedded intentional objects, but (a) this does not require any second representation, and (b) any remotely plausible picture of consciousness must have room for such objects.

⁴ Rosenthal notes this in his 1997, p. 744.

⁵ Difficulties connected with inaccurate HORs have been noted by Byrne (1997), Neander (1998), and Seager (1999). For close approximations to our worry here, see Kriegel (2003) and Raymont (2005).

⁶ If you don't like the word 'representational' here, substitute 'presentational'.

representations of C. But one cannot have *just* higher-order representations of a conscious state.

3. The source of HOR's and FOR's troubles

What leads HOR and FOR astray? We think that it is because they both hold that the following is a universal principle of representation:

RP: Representations, at least the ones relevant to consciousness, represent something other than themselves and only something other than themselves.

If RP were true, *only* HOR or FOR in some variation could be right. And if RP is not true, there would be little motive to adopt either approach.

We have just shown that RP is false. Let us capture our alternative idea that representations are self-presenting in a new principle:

SR: Representations can represent themselves as well as whatever else they may represent.

We think that we have advanced a powerful reason for preferring SR to RP. Even if our argument does not work, however, it would still be the case that SR is at least *as* plausible as RP. Since it is free of the problems facing the only approaches left open by RP, namely, FOR and HOR and, as we will see, supports an elegant, plausible model of the representational base of consciousness, we would still have reason to accept it. We turn to the model of the representational base.

4. Self-presenting representations and the representational base of consciousness

The idea that representations are self-presenting yields the following picture of the representational base of consciousness. For at least a great many representations, simply having a representation is all the representing that one needs to do to become conscious not only of what is being represented but also of the representation itself. For example, when one sees words on a computer screen, seeing them is all the representing that one needs to do to become conscious not only of the words but also of representing them – that the words are being seen (not heard, imagined, touched, etc.).

In addition, each such act of representing is all the representing needed, we believe, to become conscious of a third thing – of *who* is seeing the words, namely, oneself. We can't argue for this here but if it is so, then for each such representation, to become aware not just of what it represents and of the representation itself but also of oneself as the thing that has it, we need only that one representation.

In fact, even if a representation does not present anything other than itself, it can still represent itself and oneself as its subject. In our view, a huge number of our representations have the latter two elements of structure and most have the full tripartite structure.⁷ Our claim

⁷ . The standard picture of self-presenting representations has them presenting their object and themselves. On our picture, there can be self-presenting representations that don't have a distinct object – and self-presenting representations also present the subject of the representation.

is that representations that have at least the latter two elements of structure are the *representational base* of being conscious – of being conscious of the representation and of oneself as subject of it. Of course, when the third element of structure is present and the representation presents something other than itself and its subject, i.e., has a distinct object, it is also the representational base of consciousness of that.

We hasten to add that we are talking about the representational *base* here. We are not saying that having a representation *guarantees* being conscious of it and oneself. The view we are advancing is only that having a representation is everything *representational* needed to become conscious of having it and of oneself as its subject. Other things may be needed, too: the ability to direct attention onto one's own psychological states, for example, or the conceptual resources to go from merely representing something, say a computer, to representing it *as* a computer.

Indeed, even the requisite representation is more complicated than we have indicated so far, as we will see shortly. Before we turn to these complications, however, let us address two objections that will already have occurred to many people.

First, isn't our idea of what self-presenting representations carry information about hopelessly exotic? The more common notion of self-presenting representations is bad enough, the idea that the representational powers of these states rest upon their capacity to carry information about both themselves and their content. Surely this tripartite notion is impossible!⁸ Not at all. Something as lowly as a bar code has a tripartite information-bearing structure. A bar code contains information about what it is about, usually the item's nature and price. But it also contains information about itself – a few of the bars are an integrity check on the bar code itself. And it contains information about the thing that has it – it is physically mounted on the thing it is about. How much further the analogy holds, if any further, does not matter. In particular, whether bar-codes truly *represent* does not matter. What matters is that even something as simple as a bar code has a tripartite information-bearing structure analogous to the one that, we claim, conscious representations have.

Another analogue is one of Dretske's favourite examples, a gauge. A gauge presents information about something other than itself. For example, an altimeter contains information about the distance to the earth's surface. However, an altimeter also contains information about itself. It contains information about how far *it*, the gauge itself, is from the surface (not to mention what its dials are like, its colour, etc.). It is the gauge that contains this information about the gauge, not some higher-order gauge pointed at it. To make the analogy of the gauge complete, we would have to give the gauge one more sort of information, namely, information about the system that has it. In fact, this is easy to do. Suppose that an altimeter has to port information about altitude and itself to an avionics system. And suppose that to do so correctly, it has to recognize what sort of system it has been installed in. ('Ah, this is a Cessna Skylane.') Now we have a full tripartite information-bearing structure – and it is analogous to the one that, we claim, most mental representations have. Even something as simple as a gauge can do three distinct informational jobs all by itself, nothing else needed. (Again, whether the analogy runs any further than this structural point does not matter.)

Second, are we sure that we are dealing with a single, multiply-presenting representation here? Could there not be two or three separate representations – of the

⁸ From now on, for simplicity's sake we will just talk about the tripartite structure.

representation, of oneself, of something other than the representation? This suggestion is not plausible. The elements are very closely tied together. Take a representation of this paragraph on the computer screen. Its being a representation, its representing these words, and its representing them to oneself cannot come apart in the way that would be possible if each had its own distinct representation. Nothing could, for example, just *represent the words* and not represent them *to* anything. Nothing could just *represent to me* without representing *something* (at least itself) to me. In short, nothing could have just one of these functions without having the other. Considerations such as these suggest one representation, not two or three.

We are only halfway there but even our analysis so far has had a significant yield. We've shown that:

- There is good reason to think that FOR is false,
- The HOR approach can be reduced to absurdity and HORs are not needed for consciousness
- RP is false
- SR can be used to build a simple, unified model of the representational base of consciousness of the world and one's bodily states, of one's own representational states, and of oneself as their subject.⁹

5. Global representation, joint consciousness

So far we have talked exclusively about individual representations as understood by the tradition, roughly, representations of individual objects or small groups of objects. Such representations could certainly be self-presentational and could, perhaps, serve as the representational base of consciousness in some cases or some organisms. However, as Kant (1781/7, A110) taught us, the representations that serve as the representational base of consciousness in us are usually much 'bigger' than individual representations traditionally conceived (Brook 1994). The representations that serve as the representational base of consciousness usually have multiple objects and encompass multiple representations (as traditionally conceived). Let us call this a *global representation*.

Global representation – representing many objects as the single complex object of a single representation and many representings as aspects of this single representation.

Our points about the representational base can now be made using this notion. In our view, a global representation is all the representation needed to be conscious of its complex object, of the global representation itself, and of oneself as the 'the single common subject' (Kant, 1781/7, A350) of the representing going on in this representation. That is because a global representation represents: (i) its global object and the particular objects making it up; (ii) itself; (iii) oneself as the single common subject of one's experience.

How could a global representation be the representational base of consciousness of its complex object, itself, and its subject? A central idea here is *joint consciousness*:

⁹ A bonus: on our account, no *Ich-Vorstellung* (Frege, Husserl) is needed, either.

Joint consciousness – to be conscious of any of the objects of a global representation is to be conscious of other such objects; similarly, when one is conscious of doing any representing in a global representation, one is conscious of doing other representing involved in it.

This notion of joint consciousness applies well to consciousness of the world, i.e., to consciousness of intentional objects (including one's own bodily states) and to consciousness of representing in a global representation – in both cases, to be conscious of some items (objects, representings) is to be conscious of others. About consciousness of self, the plausible claim is slightly different. When one is conscious of oneself as the subject of one representing, one is conscious, not of a unified multiplicity, but of a single thing, oneself, as the common subject of many representings (Kant 1781/7, A350). Joint consciousness of a complex object of a global representation, of a complex representing, and/or of oneself as the subject of many representings is a central and crucial feature of our kind of consciousness.¹⁰

6. The structure of a global representation

Does a global representation contain discrete representations? Not in our view. There are good reasons to think of a global representation as not being made of 'smaller' representations. Consider a person seeing something, hearing something, and tasting something, all as parts of a single global representation in which the person is jointly conscious of the objects and the representings. How are the representings incorporated in a global representation? There are three possibilities:

1. The three acts and their objects become the object of a fourth, higher-order representation.
2. The three acts and their objects become parts of a single subsuming representation.
3. While their contents are taken up in a global representation, the three distinct representations do not survive even as parts of this state, though their objects remain.

According to (3), one's seeing something, hearing something and tasting something (at given time t) are not three distinct representations. They are three aspects of a single representation.¹¹ How might this work?

Consider what happens when I go from a state, at time t , that involves both $o1$ and $o2$ as objects, to a subsequent state that has $o1$ but not $o2$. There are two ways to represent this transition. First, we could say that where once there were two representations, $r(o1)$ and $r(o2)$, which were parts of an encompassing representational structure, $[r(o1) \& r(o2)]$, we now have only the one representation, $r(o1)$; in other words, one representation has been dropped from the representational configuration that existed at t . (1) and (2) share this picture. Alternatively, we could say that where once there was just one conscious representation, $r(o1 \& o2)$, it has

¹⁰ There are many questions about joint consciousness that we cannot take up here. Over what range of representation does it extend? Does it connect to the notion of fringe or peripheral consciousness? Is there some analogue to be found in non-conscious representation? Is there representing that is not tied to other representing in this or some analogue way? And so on.

¹¹ A similar view was outlined by James (1890, vol. 1, esp. pp. 145-61) and independently by Searle (2002) and Tye (2003), though whether Searle embraces the view is not entirely clear.

been replaced by $r(o1)$. This is the picture behind (3).

Though most philosophers adopt (1) or (2), support is seldom offered for either of them. In fact, there are good reasons to favour (3). First, (3) is representationally simpler, since it does not involve postulating representations as parts of an encompassing representation of any kind.¹² According to it, at t there was just one representation which had a complex content. The content was ‘complex’ in the sense that it had distinct contents as its parts, among which were $o1$ and $o2$. But the unified representing state does not likewise have as its parts ‘smaller’ or less complex representing states. A part-whole relation obtains among the intentional objects, but there is no parallel multiplicity of representational states.

From the perspective of cognitive function, there are other reasons to prefer (3). James argued for it this way.

Take a sentence of a dozen words, take twelve men, and to each one word. Then stand the men in a row or jam them in a bunch, and let each think of his word as intently as he will; nowhere will there be a consciousness of the whole sentence. [James, 1890, p. 160]

The conclusion he reached, using another example, consciousness of the alphabet as a whole with consciousness of each letter individually, was this:

It is safer ... to treat the consciousness of the alphabet as a twenty-seventh fact, the substitute and not the sum of the twenty-six simpler consciousnesses. [James 1909, p. 189]

Ramming representations together does not produce joint consciousness of their respective contents. Without any way of putting their contents together, we are left with a mere concatenation of representations, each member of which would be oblivious to the contents of the others. A combination of representations is not the representation of a combination. As James puts it, “Idea of a + idea of b is not identical with idea of $(a + b)$ ” (1890, p. 161). In the structure $[r(o1) \ \& \ r(o2)]$, there is no single conscious representation that takes the whole content in its scope. By contrast, in an experience of the form $r(o1 \ \& \ o2)$, all the represented objects fall within the purview of a single representation.¹³

‘But why must a global representation be *noncomposite*? Have you given any reason for holding that?’ James’ reason seems to be this: if the 27th representation, the one taking in the whole alphabet, were composite, the same problem would arise for it as arose for ramming together the original twenty-six. Here is what is behind James’ claim. However many other representations might be involved, we need at least one noncomposite representation if we are to have joint consciousness (or at minimum a joint representation) of multiple objects. Why? Because to set various objects, various words of a verse, for example, or various letters of the alphabet, *beside one another*, we must represent them together in a single representation. Being able to do this is central to our kind of cognition and consciousness. There would be no

¹² Of course, we postulate rather complicated objects of representation. However, so do (1) and (2). Indeed, so must any theory that hopes to fit the facts.

¹³ Bayne and Chalmers must have something like this in mind when they remark of a unified conscious state that “this conscious state is *not just a conjunction of conscious states*. It is also a conscious state in its own right” (2003, p. 4; emphasis added).

obvious way to do so, however, if each were contained in its own little representation.¹⁴

Moreover, once you have to posit at least one noncomposite representation of a number of objects, you no longer have any motivation to posit additional, individual representations of those objects. Anyway, (3) just is plausible, both functionally and phenomenologically (we certainly don't seem to have a representation of *just* a cat and another of *just* a mat when we see a cat on a mat). Since (2) is not an adequate at all and (3) is representationally and ontologically simpler than (1), given parsimony, if (3) accounts for the facts, there is no reason to adopt (1).¹⁵

Of course, from other explanatory perspectives, the story might be different. The individuation of representations is purpose-relative. How the brain *does* cognition might need a different story about representations, for example. But from the point of view of characterizing cognitive functions themselves, the idea that a global representation is not an assembly of constituent representations has a lot to be said for it.

Behind this discussion is one of the binding problems, the one having to do with how multiple objects get bound together in a single experience of them. Since some theorists have taken the idea of a complex noncompositional representation to require something immaterial in the mind, perhaps we should say explicitly that we have no doubt at all that producing and manipulating such representations is something that our utterly material brain does. We just don't know how yet. That empirical researchers largely ignore such representations, indeed, ignore unified representation and consciousness of multiple objects in general, is not helping.

7. Conclusion

This short paper is more a manifesto for our point of view than a detailed defence of it. All sorts of questions could be asked about the idea that representations are self-presenting, especially in our complicated tripartite way, about the idea that the representational base of consciousness is a global representation in our sense of the term, and about the idea that a global representation is not composed of other representations, even though it has a complex object.

And our attitude? The third idea first. As we have seen, there are good reasons to adopt this view. Since nobody has offered a good defence of (1) or (2), we think that these reasons are pretty strong.

That is also the way we feel about the ideas that representations are self-presenting and that global representations are the representational base of consciousness. The failure of the HOR and FOR approaches is strong support for the idea that the representations crucial for consciousness are self-presenting. And the idea that the representational base of consciousness, in beings like us anyway, is much 'bigger' and more complicated than individual representations as traditionally conceived fits our experience extremely well.

¹⁴ James urges us to reject any picture of "states of the mind which are supposed to be compound because they know many different things together" (James 1890, p. 161). Instead, for James, "we cannot mix feelings as such, though we may mix the objects we feel, and from *their* mixture get new feelings" (James 1890: 157). In fact, "We cannot even ... have two feelings in mind at once" (James 1890: 157)

¹⁵ These ideas are developed in Raymont (2003).

If sound, our picture that a global representation is not made up of other representations and is the representational base of consciousness of one's world, one's representing, and oneself would have considerable potential. Because it accounts for all three of these aspects of consciousness in the same way, it would open the way to a single, unified theory of consciousness. It would also open the way to unifying consciousness with the rest of representation and cognition. And it would open the way to nice accounts of:

- the unity of consciousness (which we get virtually for free from joint consciousness),
- consciousness of self and its special features, and,
- the subject of consciousness.

Someone should write a book about all this. In fact, someone is (Brook and Raymont forthcoming).

References

- Baars, B. (1988) *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.
- Bayne, T. and D. Chalmers (2003) What is the unity of consciousness? In: Cleersman, A., ed. 2003
- Brook, A. (1994) *Kant and the Mind*. New York: Cambridge University Press.
- Brook, A. and P. Raymont (forthcoming) *A Unified Theory of Consciousness*. Cambridge, MA: MIT Press.
- Byrne, A. (1997) Some Like It HOT: Consciousness and Higher-Order Thoughts. *Philosophical Studies* 86: 103-29.
- Cleeremans, A. ed. (2003) *The Unity of Consciousness: Binding, Integration and Dissociation*. Oxford: Oxford University Press.
- Dretske, F. (1995) *Naturalizing the Mind*. Cambridge, MA: MIT Press.
- James, W. (1890) *Principles of Psychology*. Vol. 1. London: Macmillan.
- James, W. (1909) A pluralistic universe. In *Essays in Radical Empiricism and A Pluralistic Universe*. Gloucester, MA: P. Smith, 1967.
- Kriegel, U. (2003) Consciousness as Intransitive Self-Consciousness: Two Views and an Argument. *Canadian Journal of Philosophy* 33: 103-32.
- Kant, I. (1781/7) *Critique of Pure Reason*. Paul Guyer and Allan Woods, trans. and eds. Cambridge: Cambridge University Press.
- Neander, K. (1998) The Division of Phenomenal Labor: a Problem for Representational Theories of Consciousness, in *Language, Mind, and Ontology*. Vol. 12 of *Philosophical Perspectives*, ed. James Tomberlin. Oxford: Blackwell.
- Raymont, P. (2003) The Unity and Complexity of Consciousness. *Canadian Philosophical Association* (Halifax, NS).
- Raymont, P. (2005) Rosenthal on Inaccurate HOTs. *American Philosophical Association*,

Pacific Division (San Francisco, CA).

Rosenthal, D. (1997) A Theory of Consciousness, in *The Nature of Consciousness*, N. Block, O. Flanagan and G. Güzeldere, eds. Cambridge, MA: The MIT Press.

Seager, W. (1999) *Theories of Consciousness*, chap. 3. London: Routledge.

Searle, J. (2002) Consciousness. In his *Consciousness and Language*. Cambridge: Cambridge University Press. (Originally published in *Annual Review of Neuroscience*, 23 [2000]: 557-78.)

Tye, M. (2003) *Consciousness and Persons: Unity and Identity*. Cambridge, MA: MIT Press.