

Representationalism

BENJ HELLIE

Forthcoming in T Bayne, A Cleeremans, and P Wilken eds. *Oxford companion to consciousness*.

Representationalism is the doctrine that *representation* plays a central role in *philosophical theories of consciousness*. For instance, one might say that what *makes* Jane's fleeting glimpse of a mockingbird conscious is Jane's taking a perspective on, or representing, that glimpse; or one might say that the specific *feel* of that fleeting glimpse can be understood by saying how things are from Jane's perspective, or how Jane represented things to be (it was to Jane as if a mockingbird flashed by).

Each of these views is attractive: surely we don't think that Jane's glimpse of a mockingbird could be conscious if that glimpse were entirely outside of Jane's perspective; and surely we do frequently specify the feel of a conscious mental episode by saying how things were from our perspective in that episode.

Philosophical theories of consciousness

A signal fact about psychological states is that some of them are *conscious*, or *feel* a certain way. Consider a searing pain, or a fleeting glimpse of a mockingbird: each of these feels some way. Quite plausibly, not all psychological states feel some way: cognitive science posits a series of complex computations by way of which physical stimulus causes conscious awareness, the intermediate stages of which cause consciousness without themselves being conscious.

By 'consciousness' contemporary representationalists almost invariably intend *phenomenal consciousness* [**consciousness, concepts of**], where for an episode in the mental life of a subject to be phenomenally conscious (to be an *experience*) is for there to be something the episode is like for the subject (Nagel 1974); and for an episode to be like *something* for a subject is for there to be some feature or "property" such that the episode has the feature for the subject. For instance, perhaps, an episode of experiencing pain may be *uncomfortable* for a subject. In such a case, it is said that the feature (being uncomfortable) is among the (many) *phenomenal characters* of the episode; alternatively, the episode has the feature *phenomenally* (the episode is, phenomenally, uncomfortable). So for an episode to be phenomenally conscious is for it to have some phenomenal character, to be some way phenomenally.

This elucidation of consciousness in terms of mental episodes having features *for* their subjects immediately raises two questions:

(a) When an episode is a certain way for a subject, in what further relationship between the

episode, the way, and the subject does this consist, if any?

(b) What is the possible range of the features which may be phenomenal characters?

Theories of consciousness that have been labelled “representational” have been addressed (somewhat confusingly) at answering each of these questions.

Representationality

The notion of representation employed by contemporary representationalists is part of a popular theory of what it is to take a perspective on the world, which could be called the *representational theory of perspective* [**contents of consciousness**].

A signal feature of psychological episodes is that many of them “concern”, or are “about”, or “reach toward” things outside themselves, in the sense that they are part of their subject’s perspective or take on the world. For example, a fleeting glimpse of a mockingbird is a psychological episode, in which the mockingbird is within the subject’s perspective. This *perspectivity* is a distinctive feature of psychology: when a draft through an open window scatters papers on a desk, this event does not involve anything’s perspective on anything.

The notion of perspective is somewhat metaphorical; the representational theory of perspective explicates it in two stages. First, each aspect of a subject’s perspective is understood as a taking of an attitude toward a way the world could be. Consider a judgement, or a hope, that Portugal wins World Cup 2010. These are both, in a sense, aspects of a subject’s perspective on the world. Both fit the mold of being attitudes toward a way the world could be: that Portugal wins World Cup 2010 is a way the world could be---if Portugal *does* win World Cup 2010, then the world *is* that way, if it *doesn’t*, then the world *isn’t*; the *judgement* involves taking the attitude of *judging* toward this way the world could be, while the *hope* involves taking the attitude of *hoping* toward this way.

Second, the notion of taking some attitude or other toward a way the world could be is explained in terms of a *norm of representational correctness*. Which way-the-world-could-be a mental episode is an attitude toward is given by how the world would have to be for that episode to be correct as a representation of the world. A judgement or hope that Portugal wins World Cup 2010 is correct as a representation of the world if Portugal does win World Cup 2010; incorrect as a representation of the world if Portugal does not win World Cup 2010. Correctness as a representation of the world is, like correctness as a use of a salad fork, a “normative” feature, resulting from association with some system of value: in particular, from the system of values associated with representations as such.

For a mental episode to be an aspect of a subject’s perspective, on the representational theory of perspective, then, is for it to be *representational*: to be subject to the system of values associated with representations as such. Say that a property like *being correct as a*

representation just in case such and such is a representational property.

Either stage of the explanation of perspective in terms of representational values may be questioned. Particularly relevant to the study of consciousness are challenges to the first stage: some philosophical theories of consciousness accord a central explanatory role to perspective, but reject the representational theory of perspective [**acquaintance, intentionality**].

Higher-order and first-order representationalism

Since the theory of consciousness aims to answer both questions (a) and (b), representation may play a role in answering either question. Indeed, philosophers have appealed to representation in answering both questions.

A theory which appeals to representation in answering question (a) (Rosenthal 2002; Kriegel and Williford 2006; Lycan 1996) would take the following general form: for a mental episode to be a certain way for its subject is just for the episode to be represented to the subject as being that way: for the subject to undergo some mental episode which is correct as a representation if, and only if, the episode is that way. Such an answer is sometimes referred to as *higher-order representationalism*.

Alternatives to higher-order representationalism are legion. For instance, there is the view that for a mental episode to be a certain way for its subject is: for it just to *be* that way (assuming, of course, that the feature is a possible phenomenal character in accord with the answer to (b)); or for the episode's being that way to be in a position to play a central role in its subject's cognition (for it to be "poised": Tye 2000); or for the episode's particular instance of being that way to have a special conscious "glow" which cannot be understood in more basic terms; or for its being that way to be within the subject's perspective in some way not compatible with the representational theory of perspective.

A theory which appeals to representation in answering question (b) (Siewert 1998; Tye 2000; Byrne 2001; Chalmers 2005), would take something like the following form: a feature may be a phenomenal character only if it is a representational property. Such an answer is sometimes referred to as *first-order representationalism*.

Alternatives to first-order representationalism are also legion. For instance, there is the view that *any* feature which a mental episode can have can be a phenomenal character, assuming it and the episode together meet the condition mandated in the answer to (a); various familiar physicalist and functionalist answers---e.g., that the phenomenal characters are certain special brain features [**functionalist accounts of consciousness**]; the view that a property is a possible phenomenal character only if that property is of the special conscious type, a type which cannot be understood in more basic terms; and the view, once again, that some nonrepresentational properties characterizing a subject's perspective are phenomenal characters.

First- and higher-order representationalism are compatible. A view answering both questions is that every phenomenal character is a representational property, but that not every representational property is a phenomenal character---to be “promoted” to phenomenality, such a property must be itself represented (Lycan 1996).

Master cases for representationalism

The master case for higher-order representationalism appeals to our inability to make sense of an episode’s being some way for its subject despite its being *no way at all* from the subject’s perspective---despite the subject’s being *utterly blind* to the episode (Kriegel and Williford 2006). If this really is incoherent, then if an episode is some way for its subject, it is some way from the subject’s perspective. And, granting the representational theory of perspective, if an episode is some way for its subject, it is represented to its subject as being that way. It would then be natural to explain this by holding that what *makes* the episode be some way for its subject is that it is represented to its subject as being that way. That is what the higher-order representationalist says in answer to (a).

The master case for first-order representationalism stems from introspective or “phenomenological” study of a variety of phenomenally conscious episodes: to fully characterize what it is like for me to see now I must say that from my perspective, a yellow architect’s lamp jitters on its mount (Siewert 1998); to fully characterize my conscious emotions in November 2004 I must mention dejection at election results and concern about the future (Tye 2000)---both aspects of my perspective. (This phenomenological point is one of several made in an attempt to elucidate G. E. Moore’s famous discussion of a “transparency” somehow involved in reflection on experience: Hellie 2006, forthcoming b [**transparency**].) If this is correct, and if it is granted that the features revealed to phenomenological study are phenomenal characters, this makes a strong case that at least many phenomenal characters of the most important mental states are properties concerning how things are from one’s perspective. First-order representationalism follows if we plug in the representational theory of perspective.

Representationalism and reduction

Ambitious projects for the reduction of consciousness to something more basic have been pursued in the hope of locating the place of this richly colorful and flavorful feature of the world in the austere mathematical world of physics [**physicalism; epiphenomenalism; the explanatory gap; the hard problem**]. A number of partial successes in reducing certain varieties of representation to something a bit more austere have suggested the use of representation as a way-station: first, reduce consciousness to some kind of representation; then, reduce that kind of representation to something austere (Tye 2000).

The two-stage reductive strategy faces the difficulty that, in combination, the fact that consciousness seems hard to reduce to the physical and the fact that consciousness seems

easy to reduce to a certain variety of representation undermine the view that the variety of representation to which consciousness reduces is the same as that which reduces to physics---a problem especially pressing if representation is to be understood as is standard as involving a variety of *value*. If the cognitive scientists are correct, not all representation brings about phenomenal character; so perhaps the sort which does is an irreducible sort of *conscious* representation (Stoljar forthcoming).

We now turn to some local issues of detail in the development of higher-order and first-order representationalism.

Issues for the higher-order representationalist

Much of the literature on higher-order representationalism is devoted to peculiarities of three of its variants: the *higher-order thought* (Rosenthal 2002), *inner sense* (Lycan 1996), and *Brentanian* (Kriegel and Williford 2006) theories. These views differ along two dimensions. The first concerns the “topology” of the “inner” representation of the episode. On the first two views, this inner representation is other-representation: the representing episode is distinct from the represented episode; while on the third, it is self-representation: the representing episode *just is* the represented episode. The second of these dimensions concerns the taken toward the inner representation. On the higher-order thought view, the attitude is more like judgement; while on the inner sense view, the attitude is more like perception.

Concerning the latter dimension, it’s clear that inner representation is not exactly like ordinary cases of judgement or perception. Against exact resemblance with judgment, babies arguably have experiences without being sophisticated enough to form judgments about them; also, the phenomenal characters of our perceptual experiences seem limitlessly intricate while our judgements seem comparatively crude. Against exact resemblance with perception, whenever one perceives a mockingbird, one has an experience that is distinct from the mockingbird itself; but a conscious perceptual experience of a mockingbird does not seem to involve both a perceptual experience of the mockingbird and a *distinct* perception *of the perception* of the mockingbird.

Concerning the former dimension, the motivation for higher-order representationalism extends naturally to the self-representational view. If an episode’s being like something for its subject requires that episode’s being within the subject’s perspective, then shouldn’t that perspective also be like something for its subject? Otherwise, it would seem, the whole arrangement would merely take nonconscious representation of a mockingbird, and, as it were, push it inward, resulting in nonconscious representation of representation of a mockingbird---without any additional gain of consciousness. But if so, this would seem to set off an explosion of inner representation without bound. Such an explosion would involve either an infinite chain of distinct other-representing mental episodes, or at some point a circle, in which some episode represents itself. Since, quite plausibly, we have only finitely many mental episodes at once, the former is right out. Hence, consciousness

requires, at some stage, self-representation; and why delay the inevitable?

A central question for the higher-order representationalist is whether, always, if an experience is a certain way phenomenally, it is that way in fact. Could it be that, while an experience is unpleasant for its subject, it is *in fact* pleasant, or at least neutral---or that, while an experience is, for its subject, like seeing a green square, it is *in fact* like seeing a red octagon? The idea beggars comprehension. And yet the higher-order representationalist seems pressured to accept this possibility. After all, *outer* representation is fallible: sometimes green squares look like red octagons. Inner representation would certainly be peculiar if it were utterly immune to error.

Issues for the first-order representationalist

The literature on first-order representationalism contains extensive discussion of the following three issues.

First, consider the phenomenal contrast between seeing and feeling an X, despite both representing that something X-shaped is before one. The first-order representationalist can explain this difference by appeal to *manners* of representing, via the “impure” representational properties (Chalmers 2005) of *visually* versus *tactually* representing that something X-shaped is before one. Or consider blurrily seeing an X: first-person reflection arguably reveals the experience to be blurry, but blurriness is also plausibly a manner of representing (Crane forthcoming). The same strategy could perhaps distinguish *perceiving* from *judging* that something X-shaped is before one [**phenomenology of abstract thought**]. Such manners influence how an experience is for its subject, seem to be qualities of the experience (just as a slow run is an episode with slowness as a quality), and fail to be representational properties; so accepting impure representational properties as phenomenal characters requires accepting nonrepresentational phenomenal characters. This involves some departure from the most full-blooded first-order representationalism. Still, these features seem to qualify and depend on the representational features (as the slowness of a run depends on its being a run), so that representational properties still end up fundamental to consciousness.

Second, for some phenomenal characters, the core claim of the phenomenological case for first-order representationalism is not especially contentious: for instance, when one sees an X before one, arguably, first-person reflection on the experience reveals it to be successful as an experience if and only if representationally correct if and only if something X-shaped is before one. (Still, this is contested: perhaps first-person reflection reveals only a nonrepresentational visual “taking in” of or “acquaintance with” an X (Travis 2004) [**acquaintance**].) By contrast, for other phenomenal characters, the representational theory of perspective does not seem to be a natural fit. A visual experience in which nothing even seems to be seen---such as a visual experience with closed eyes---does not clearly involve any representation of anything; but it may seem to involve a non-representational perspective on something---a field suffused with a sort of “blackness” (Hellie 2006).

Third, the “inverted spectrum” [**the inverted spectrum**] threatens first-order representationalism. Very briefly, the issue is this: (i) plausibly, there could be a subject Abnorm who has the same phenomenal character when seeing green as the rest of us have when seeing red, and vice versa. This need not interfere with Abnorm’s ability to navigate his environment. But if not, we should assume Abnorm correctly represents the colors of things. But then (ii) there need be no difference between the representational properties of the experiences of Abnorm and of a normal subject when seeing certain scene, despite the evident difference in their phenomenal characters. But then it seems that phenomenal character varies independently of representational properties, against first-order representationalism. The literature contains a very extensive discussion of attempts to block (i) (Tye 2000; Lycan 2001) and (ii) (Chalmers 2005; Shoemaker 2006).

References

- Byrne A (2001). Intentionalism defended. *The Philosophical Review*, **110** 199--240.
- Chalmers DJ (2005). The representational character of experience. In B Leiter, ed. *The future for philosophy*. Oxford University Press, Oxford.
- Crane T (forthcoming). Intentionalism. To appear in A Beckerman and B McLaughlin, eds. *Oxford handbook to the philosophy of mind*. Oxford University Press, Oxford.
- Hellie B (2006). Beyond phenomenal naivete. *The Philosophers' Imprint*, **6/2**.
www.philosophersimprint.org/006002.
- Hellie B (forthcoming a). Higher-order intentionalism and higher-order acquaintance. To appear in *Philosophical Studies*.
- Hellie B (forthcoming b). That which makes the sensation of blue a mental fact. To appear in *European Journal of Philosophy*.
- Kriegel U and K Williford, eds. (2006). *Self-representational approaches to consciousness*. MIT Press, Cambridge, MA.
- Lycan WG (1996). *Consciousness and experience*. MIT Press, Cambridge, MA.
- Lycan WG (2001). The case for phenomenal externalism. *Philosophical Perspectives*, **15** 17--36.
- Rosenthal DM (2002). Explaining consciousness. In DJ Chalmers, ed. *Philosophy of mind: classical and contemporary readings*. Oxford University Press, Oxford.
- Shoemaker S (2006). On the ways things appear. In TS Gendler and J Hawthorne, eds. *Perceptual experience*. Oxford University Press, Oxford.
- Siewert C (1998). *The significance of consciousness*. Princeton University Press, Princeton.
- Stoljar D (forthcoming). Consequences of intentionalism. To appear in *Philosophical*

Studies.

Travis C (2004). The silence of the senses. *Mind*, **113** 57--94.

Tye M (2000). *Consciousness, color, and content*. MIT Press, Cambridge, MA.