

Explaining Intentionality: replies to Critics

Tim Crane

I am very grateful to the authors of these articles for engaging with my paper on the explanation of intentionality. I have learned a lot from their comments and criticisms. I see some broad themes in common among the criticisms, so rather than respond to each author individually I will attempt to unify my responses around three themes: reductionism about intentionality, the metaphysics of intentionality, and the methodology of its investigation.

1. Understanding intentionality: reduction, physicalism and dualism (Bordini, Fernandez, Janssen-Lauret, Ivanov, Kingsbury, Rinner, Chen)

The overall point of the target article is to criticise the idea that any theory of intentionality must answer the ‘question of aboutness’: *what makes it that case that any mental state can represent anything at all?* I took the question as usually coming with two further conditions: first, that the answer to this question must be wholly general; and second, that the answer must not use any intentional or representational notions. Let’s call these the ‘generality condition’ and the ‘non-intentional condition’.

I used Hilary Putnam’s famous argument at the beginning of *Reason, Truth and History* as representative of this way of conceiving of the problem of intentionality. I did not say, of course, that it is the *only* way in which the problem has been conceived, nor the only possible way — just that it is one very common way.

In the argument which follows his famous thought-experiment about the ant, Putnam’s makes the crucial assumption that what applies to pictorial and linguistic representation applies to mental representation (or thought) too. I identified two ways in which this assumption might be defended: either by appealing to physicalism, or by appealing to what I call anti-psychologism. I argued that the appeal to physicalism makes Putnam’s argument redundant, and (relying on some other work of mine) that anti-psychologism is implausible and not mandatory. I then offered an alternative way of thinking of the place of intentionality in the natural world that does not impose the generality or non-intentional conditions.

Again, I do not say that physicalism and anti-psychologism are the *only* reasons why someone might defend Putnam’s crucial assumption. Davide Bordini argues plausibly that one might defend this assumption without these reasons, but on broadly methodological grounds. In effect, Bordini proposes the generality condition as a methodological constraint, which could lead to Putnam’s assumption without assuming physicalism or anti-psychologism. Jordi Fernandez makes a similar point.

I agree someone could do this; my claim was that physicalism and anti-psychologism are individually sufficient to motivate Putnam's assumption. I did not say they were jointly or even disjunctively necessary. Bordini's historical examples from Locke and Descartes are illuminating, though I wonder how Descartes's talk about 'images in the brain' is consistent with his official view that conscious intentional episodes are modes of the attribute of thought. I need to consider the historical picture more deeply, and I am grateful to Bordini for drawing my attention to these points.

A number of my critics treat the paper as proposing a non-reductive theory of intentionality, and in a sense I am. I certainly reject the idea that one must pursue the task of *defining* intentionality in reductive terms, where this means defining it without mentioning any intentional or representational notions. (I am therefore declining to answer what Jordi Fernandez calls 'the definitional question'.) But this is not the only thing that 'reduction' has meant in contemporary philosophy.

In earlier work (Crane 2001, chapter 2), I distinguish between ontological reduction and explanatory reduction. (I do not claim, of course, that I invented this basic distinction; I just hope to have clarified it.) In general I reject ontological reduction of mental entities — e.g. the identity theory — on the grounds that no persuasive arguments have been provided for it. A properly naturalistic philosophy, I say, can accept the reality of the psychological for what it is, and accept psychology and cognitive neuroscience as the sciences of this reality. These sciences do not need ontological reduction to validate them.

But in the target paper, I am neutral on ontological reduction. My aim in the paper is to identify the network of ideas and assumptions underpinning much contemporary philosophy of intentionality, and to challenge them. It would be possible to challenge the generality and non-intentional assumptions while maintaining ontological reduction (or physicalism, in one of its many senses), as Bordini and Stefan Rinner point out. Or it would be possible to challenge these assumptions and reject ontological reduction (or physicalism), as Frédérique Janssen-Lauret does. When it comes to these assumptions about intentionality, physicalism is a wheel that might turn, but is no part of the mechanism.

So the main point of the paper is neutral on the question of physicalism; nowhere do I endorse or deny physicalism. I am puzzled therefore why Dimitri Ivanov says that I am offering a physicalist account of intentionality. Nowhere in the paper do I say I am offering a physicalist account. Ivanov is under no obligation to consult my other works; but if he had, he would have found me rejecting both physicalism and ontological reduction. So his attribution of a physicalist theory to me is something of a mystery.

On the other hand, I offer no argument in principle against *explanatory* reduction, and I don't think, as a naturalistic philosopher, that we should object to explanatory reduction on *a priori* grounds. An explanatory reduction is an explanation of a phenomenon or theory in

terms of a 'lower-level' phenomenon or theory — e.g. Ernest Nagel's classic example of the explanation of thermodynamics in terms of mechanics. If such reductions are available, then as explanations they advance our knowledge, so no-one should reject them *a priori*. If such reductions are not available, then we should look for other kinds of explanations of those things we want to understand.

Kuei-Chen Chen claims that I reject the kind of explanation offered in cognitive science, but I am not sure why he thinks this. The approach I adopt in the final section of the paper involves first identifying mental capacities, and then proposing empirical explanations of how these capacities work. Explanations in terms of sub-personal mental representations belong here. A no point in this paper, nor elsewhere in my work, do I reject the possibility of such explanatory reductions.

Chen is right, though, that I reject reductive *definitions* of intentionality and related notions. I see no future (and no point) in giving an account of intentionality without mentioning intentional notions. There is nothing in cognitive science which requires that intentional notions should be eliminated by such a definition. Explanatory reduction, as I understand it, is not elimination by definition.

Chen correctly conjectures that I will insist that the idea of information *as used in cognitive science* is a representational notion, and asks rhetorically whether *conditional probability* and *physical magnitude* are representational notions.

My answer is that of course the latter notion is not representational; but no-one thinks that cognitive science can make do only with physical concepts like the concept of a magnitude. Conditional probability, on the other hand, must be understood in terms of the notions of truth and proposition (after all, probability is the probability of the truth of a proposition) and is therefore representational. The fact that on a view like Nick Shea's, 'representations of the sort central to the cognitive sciences can be described "in non-semantic, non-mental, non-normative terms"' is not really to the point — the point is rather whether these semantic and intentional descriptions can be defined away and dispensed with, not whether they also have non-semantic and non-normative descriptions.

I should have made it explicit, perhaps, that explanatory reduction is perfectly compatible with ontological non-physicalism. It should be obvious, though, that explanatory reduction is also perfectly compatible with ontological reduction, and therefore with physicalism. In his commentary, Stefan Rinner in effect underlines this point by pointing out that 'physicalism is compatible with the claim that (at least some) mental representations represent intrinsically'. Rinner argues that Recanati's theory, which takes Fregean modes of representation to be mental files, not only entails that mental files represent intrinsically, but also that they are physically realised in the brains of the agents. He therefore concludes that 'the example of Recanati's theory of mental files .. shows that, contrary to Crane, the existence of intrinsic mental representation is compatible with physicalism'.

Given the wide variety of things that 'physicalism' can mean, I would not have said that intrinsic intentionality was incompatible with physicalism. I see no reason for denying that intrinsic intentionality is compatible with global supervenience versions of physicalism, for example. My point in the paper is only that physicalism could be a motivation for the claim that there must be 'thought forms' in Putnam's sense; but in this case a physicalist would not need Putnam's argument to explain why there is a problem of intentionality, since they could just see the problem as a request for an explanatory reduction of intentionality which is consistent with physicalism. Rinner's appeal to Recanati's theory in effect takes this approach: Recanati's theory offers an explanatory reduction of intentionality in terms of ideas like mental files.

Whether we should take a 'reductionist' or a 'non-reductionist' approach to intentionality is therefore a complex question. Let me therefore sum up the approach to reduction taken in the target paper.

My overall aim there, as explained above, is to reject the generality condition and the non-intentional condition as requirements for an explanation of intentionality. Rejecting the non-intentional condition is the same as rejecting the need for a definitional reduction (i.e. the need to answer Fernandez's 'definitional question'). In the final section of the paper, I endorse the prospect of explanatory reduction, so long as it does not require conformity to the non-intentional condition.

When it comes to ontological reduction or physicalism, the paper is neutral. Physicalism is consistent with the rejection of the generality condition and the non-intentional condition as requirements for an explanation of intentionality, as a number of the commentators point out. However, in other work I have been sceptical of physicalism both as a doctrine, and as a question worth addressing.

This brings me to dualism. While Ivanov says that I offer a physicalist account, Frédérique Janssen-Lauret claims I am a dualist. I can only hope that these divergent interpretations are due to the fundamental unclarity of these categories rather than the unclarity of my writing.

In the paper I did mention Cartesian dualism, but only as an analogy to make a point about Putnam's notion of 'thought forms'. My point was that simply asking 'how can thought have representational power?', without the assumption of 'thought forms', makes as much sense as asking 'how can a mental substance think?' on a Cartesian view of mental substance.

The fact is that in the target paper I made a commitment neither to Cartesian dualism nor to any kind of physicalism. Of course, if physicalism is by definition simply the denial of Cartesian dualism, then I must concede that we should all be physicalists. But physicalism

was surely supposed to say more than just this! My own view is that if we don't accept an ontology of substance in the Cartesian sense, then we are not compelled to choose between physicalism and dualism, but accept mental things as they are, among the many kinds of things we encounter in the natural world.

I totally agree therefore with Janssen-Lauret's claim that 'naturalism is perfectly compatible with the denial of physicalism'. I would be also happy to endorse her naturalistic dualism, since this doctrine is essentially the acceptance of mental entities within a naturalistic framework, plus the rejection of physicalism, understood here as the extreme and implausible view that the only things there are are those 'posited by the science of physics'.

Since mental entities are not posited by the science of physics, and yet they manifestly exist, then physicalism in Janssen-Lauret's sense is quickly refuted. She rejects other, weaker, forms of physicalism by arguing that the metaphysical notions involved (supervenience, grounding etc.) are physicalistically unacceptable. I myself have little sympathy with the notion of 'grounding'. If 'grounding' is just an alternative word for explanation, then we already have enough words. And supervenience is too metaphysically empty to do the work physicalism needs. So I am sympathetic to Janssen-Lauret's claims here. I mentioned the weaker versions of physicalism in my paper only because nothing turned, in the argument of the paper, on which specific conception of physicalism was under discussion.

The dualistic component of Janssen-Lauret's naturalistic dualism is, to my mind, very plausible and should not be controversial. I do not share her enthusiasm for Quinean approaches to ontology: that is, regimenting discourse according to a preferred vocabulary, and taking quantification as the expression of ontological commitment. But this would be a discussion for another occasion.

What about the 'naturalism' in naturalistic dualism? Like physicalism, naturalism has meant many things, but I would agree with Janssen-Lauret that the term is best reserved for a methodological or epistemological doctrine (see the discussion of explanatory reduction above). Again, I would not count myself as a Quinean naturalist. But I took myself to be proposing a naturalistic approach to intentionality in the final part of the paper.

My view is a kind of dualism in the undeniable sense Janssen-Lauret proposes, but not in any more demanding sense; and it isn't physicalism. One of my ambitions is to try and transcend or deconstruct the apparent forced choice between physicalism and dualism, as based on categories that are fundamentally unclear or implausibly demanding. For this reason, I would rather avoid calling myself a naturalistic dualist, since it gives the impression of accepting the terms of a debate I would rather deconstruct. But if others want to call me this — so long as it is understood in Janssen-Lauret's sense — then I will not object.

2. Understanding intentionality: the metaphysics (Bordini, Gow, Ivanov)

Another issue on which I differ fundamentally from some of my critics is how we should conceive of the metaphysics of intentionality itself. Should it be understood as a relation (like acquaintance, as Ivanov argues) or is it a non-relational 'pointing' which falls short of relationality? Is the concept of intentionality the concept of a relation or of an intrinsic property?

Gow argues that there is not a real choice here, because non-relational intentionality and the relation of representation are simply two different phenomena. The relation of representation, she claims, is amenable to pluralistic explanation, but non-relational intentionality is the better candidate for being the mark of the mental. In the same ballpark, Bordini makes a sympathetic to Uriah Kriegel's distinction between subjective and objective representation, where objective representation might be amenable to physicalist reduction and subjective representation not.

I take Gow's distinction to be a stipulation about how to use of the words 'intentionality' and 'representation', rather than a theory of what these words mean in in the philosophical tradition (intentionality) or their ordinary usage (representation). After all, in its philosophical use, the phrase 'relational intentionality' is not a contradiction in terms; and there is nothing about the meaning of the word 'represent' which prohibits us from saying that you can represent what does not exist. Given this stipulation, Gow develops an interesting account of what should be treated as relational and what is non-relational (or as she calls it, rather misleadingly in my view, 'adverbial' — adverbs are not really the point, I feel). The account deserves more discussion, as does Kriegel's distinction.

Gow in effect raises a challenge about how the same feature (intentionality) can be both a relation and not a relation. On the one hand, it is natural to say that thinking about X could be the same kind of thing whether or not X exists (even those who believe in object-dependent singular thought think that this is true for some kinds of thought). But on the other hand, when X does exist, what is wrong with saying that in this case one's thought consists in a relation to X? Having struggled with this dilemma myself, I understand the urge to try and divide the concepts in a way that corresponds to the different phenomena.

However, I now think that the best way to resolve this is to reject the idea that there is such a thing as the property (or relation) of intentionality at all. This is the radical message of the last part of my paper. Just as there is no such thing as the real relation of parthood, so there is no such thing as the real property or relation of intentionality. It remains true that things have parts and we can say general things about parthood. But once we have said everything about how the parts of specific types of objects come together to form whole objects of those types (as it may be, organisms or automobiles) there is no further question of whether the real relation of composition holds between these parts, and thus

no obligation to answer what Peter van Inwagen has called the ‘special composition question’, or other such questions of mereology.

I do not pretend this approach to parts and wholes is uncontroversial or obvious, only that that it is an analogy for what I want to say about intentionality. The analogy is this. We can describe what we mean by intentionality, what general range of phenomena this is meant to capture, and what features may or may not be shared by all these phenomena. But this does not mean that there is a real or natural property (or relation) which these descriptions pick out. Once you have given the empirical descriptions of how the mechanisms of the different intentional capacities work, then we should not need to ask about the metaphysical nature of the distinct and real relation or property of intentionality. I propose we have no use for such a property or relation.

Nor would I have any use for philosophical concepts which attempt to ground intentionality in some fundamental psychological relation like acquaintance, as proposed by Ivanov. I have criticised the philosophical notion of acquaintance elsewhere, so I will not repeat these points here (Crane 2012).

3. Understanding intentionality: the methodology (Bordini, Cheng, Coehlo Mollo, Fernandez, Kingsbury, Woodling)

The third theme in these commentaries is how we should go about understanding intentionality — what strategies and methods should we use. My paper offered what I took to be a naturalistic approach, but a naturalistic approach that was not committed to the non-intentional and generality conditions.

The naturalistic approach sketched there is obviously programmatic and leaves many questions unanswered. Something I do not touch on is how to relate intentionality to the semantics and pragmatics of intentional attributions. One tradition in analytic philosophy has been to study the semantics of attitude attributions first, and then draw conclusions about the structure of the attitudes from this study. My general approach, which I label ‘psychologistic’ (Crane 2014), would take the reverse approach. However, a full account of the phenomena also would need an account of attitude attributions too.

Here I am grateful to Casey Woodling for his helpful account of how Mark Sainsbury’s ‘display theory’ of the attitudes fits together well with my approach to intentional phenomena. Woodling carefully explains too how Sainsbury could adopt the conception of intentional objects I defend in *The Objects of Thought* (2013), despite the fact that Sainsbury himself rejects it. Woodling shows that *Intentional Object* is not an ontological category, and cannot be understood independently of understanding the contents of intentional states. This offers us a partial understanding of how intentionality can fit into the natural world. I have nothing to add here to Woodling’s fine discussion.

Some of my commentators argue that the generality and non-intentional conditions need not necessarily stand or fall together. Justine Kingsbury points out that the two conditions might be in tension, and that the physicalist who accepts the non-intentional condition would be well-advised not to accept the generality condition. That seems to me correct.

Jordi Fernandez argues that defining intentionality (in non-intentional terms) should be general, but accounting for it empirically ('productively' in his sense) can be pluralistic. He seems to think that it would be a relatively simple matter to give a non-intentional definition. So the task of explaining intentionality is too easy, he suggests, if it is restricted to giving such a definition; but too hard if the generality and non-intentional conditions are imposed on the empirical project of explaining intentionality. He concludes that there is no need to impose these conditions on a genuinely empirical explanation of intentionality.

Of course I agree with the latter point; my disagreement with Fernandez here is over whether my argument's target is a straw person. He says he is inclined to think that:

those naturalistic philosophers who approach the project of explaining intentionality by assuming that they need to address the question of aboutness never expected that they would need to clear such a high bar [of assuming the generality and non-intentional conditions] in the first place.

However, he provides no evidence for this inclination. In the paper, I do give some evidence and the whole intentionality industry from the 1980s and 90s provides a vast amount more. A few dissenters like Daniel Dennett and Frances Egan rejected the general picture, but most of the intentionality project was all about finding the essence of representation, characterised in non-representational terms.

My other substantive disagreement with Fernandez is about whether there can actually be an adequate general definition of what it is to represent something, stated in non-intentional terms. Fernandez says that the definition should be general because 'we are looking for a definition, or analysis, of what is to represent something'. I can agree that we can say general things about what it is for 'mental states, fuel gauges and linguistic utterances' to represent: that they all have objects, that they can be correct or incorrect, that they can convey information about the environment, for example. These are some general claims we can make about intentionality or representation. Fernandez implies that such claims will be 'uninformative'. I disagree. Saying what is meant by a mental state 'having an object' is a significant task, even if one can help oneself to intentional notions. My own efforts to answer this question (Crane 2013) may be mistaken, but to be mistaken a claim must be aiming to be informative.

Fernandez also says that to develop this into a definition we should use only non-intentional notions because this will reduce the notion of intentionality 'to other notions which are easier for us to understand'. The recent history of the philosophical discussion of

intentionality tells a different story. I challenge anyone to say that asymmetric dependence (Fodor), the mathematical theory of information (Dretske) and proper function (Millikan) are — for all their importance — ideas which are easier for us to understand than the idea of thinking about an object.

Like Fernandez, Dimitri Coelho Mollo is another who is sceptical that the question of aboutness (as I conceive it) is widely accepted as the starting point for a theory of intentionality. He cites Nick Shea and Karen Neander as philosophers who do not accept its generality assumption. Kingsbury points out too that Neander rejects the generality assumption. In response I would point out that both these philosophers propose broadly teleological theories of content. What is the point of these theories? I think they are proposed as answers to the general question ‘how is representation possible in a physical world?’ Shea’s approach starts with the biological functions of simple sub-personal representations, but it is no less an answer to that question, for all that.

However, I am happy to concede that Kingsbury and Coelho Mollo may be right that the generality and non-intentional conditions are not imposed by all philosophers today on an account of intentionality; and I certainly agree with Fernandez and Kingsbury that they should not be so imposed on any such account. The more that is done to remove the insistence on these conditions, the better.

Finally, Tony Cheng says that I have overlooked or ignored the fundamental requirement on a philosophical account of intentionality, which is that it should give a ‘transcendental explanation’ of the property of intentionality.

Cheng says that we all knew that ‘perception, memory, language, reasoning, sensation are topics in cognitive psychology’ but ‘there are philosophically distinct reasons why philosophy still has something significant to say about these topics in an age of science: intentionality is one such property that sciences typically do not deal with’. He then says that my failing to notice this fact is what leads me to my view of the explanation of intentionality (which for some reason he calls ‘Quinean’: for the record, I am not in any sense a Quinean).

Here it is Cheng who misses the point. I do not ignore the transcendental question. Rather, I reject it. Cheng does not tell us what the ‘philosophically distinct reasons’ are for raising the question; when he does this, I would be happy to discuss them. But in the meantime, I shall conclude here by repeating what I said in reply to Gow and Ivanov.

We will make progress here by not asking ‘what is intentionality itself?’ but rather by looking at the actual empirical modes of intentionality and their concrete nature or mechanisms. Of course, as I said in reply to Fernandez and Gow, we can talk about intentionality as the ‘mind’s direction on its objects’ and so on. This talk is illuminating for identifying the general features of anything that counts as a mental (that is, intentional)

capacity at all. But I need more argument to be persuaded that we need to give a transcendental or 'grounding' explanation of this talk.

References

Tim Crane (2001) *Elements of Mind* (Oxford: Oxford University Press)

_____ (2012) 'Tye on Acquaintance and the Problems of Consciousness' *Philosophy and Phenomenological Research* 84: 190-98

_____ (2013) *The Objects of Thought* (Oxford: OUP)

_____ (2014) *Aspects of Psychologism* (Cambridge, MA: Harvard University Press)