

Replies

Crispin Wright

I am grateful to each of my commentators for their perceptive and interesting observations, and to the Editor for giving me an opportunity to reply in print.

SAINSBURY

In my paper, I offered a “worst-case” diagnosis of the McKinsey argument as involving a failure of transmission of non-empirical warrant from its premises to its conclusion (worst case because it allows, what seems to me uncertain, that a sufficiently strong externalist major premise can indeed be warranted non-empirically). Mark Sainsbury is dissatisfied with this diagnosis. He believes that the real failure of both the McKinsey and the Putnam arguments resides rather in a question-begging use of principles of disquotation. I’ll say briefly why I’m not persuaded by that diagnosis. But first I want to react to the, as it seems to me, rather heavy weather which Sainsbury makes of understanding the notion of transmission of warrant in the first place.

Sainsbury glosses my initial characterisation of what it is for an argument to transmit a warrant from its premises to its conclusion as a conditional: "if one acquires warrant for its premises and one recognises its validity, then one acquires warrant for its conclusion". So, he observes, "transmission should fail, presumably, if the antecedent of the conditional is true and the consequent is not". He then remarks that this account of transmission failure will not engage in a case where, owing to the operation of some strong form of externalism, the premise in question is actually content-defective. For in that case one will not have acquired any warrant for the premise in the first place. He concludes "we therefore need to get a better fix on how failure of transmission is to be demonstrated".¹

No doubt we do need such a "better fix"—Sainsbury here focuses on something that was only put forward as a first approximation. But his reaction is nevertheless uncomfortably suggestive of a muddle. There wasn't supposed to be a transmission failure in the case where there actually *is* an illusion of content—that is indeed, to be sure, a case where there is no warrant for the premise in the McKinsey argument. The issues about transmission arise only when we take it that *there are indeed warrants* for the premises of the argument and they focus on the rational presuppositions of those putative warrants. A case of transmission failure is a case where in order to make out that one is warranted in accepting a premise of the argument in question, one would need to establish some kind of warrant—perhaps merely an "entitlement"—for the conclusion of the argument independently of its being the conclusion. That's the common structure exhibited both by cases of simple question-begging and by examples of the two templates—respectively featuring information-dependent warrants and non-inferential warrants—which I distinguished. Counterexamples to transmission are thus exactly what Sainsbury's conditional formulation would suggest they are: cases where someone may acquire warrant for the premises of an argument, and recognise the argument's validity, and yet not thereby acquire warrant for the conclusion precisely because it was only because they already had warrant for the conclusion that they were able to acquire a warrant for the premises.

Sainsbury's discussion of the zebra/painted mules example suggests that he not merely rejects the diagnosis of transmission failure in the McKinsey case, but is inclined—notwithstanding some more concessive-sounding earlier remarks—to reject the notion of transmission failure altogether. He writes

It seems that the casual observation that warrants the belief in zebras is not sufficient to repudiate a serious *assertion* to the effect that the animals are mules. This is consistent with supposing that, in the absence of such an assertion in the context, the subject is in fact warranted in believing that the animals before him are not mules: if...he were to come to believe that the animals are not mules, that belief would be warranted. The warrant would consist in the observations that warranted the belief that the animals are zebras, along with whatever warrant he has for thinking that zebras are not mules...

The zebra example effectively revolves around whether one could refute, merely on the basis of casual observation, one who *claimed* that some observable animals were not zebras but painted mules; it seems clear that one could not. In this case...we seem not to have a counterexample to transmission... It is not that we have a warrant which fails to cross [sic] the rails of entailment but rather that, in the envisaged circumstances, we do not have an undefeated warrant at all.²

Here Sainsbury is suggesting that the impression of transmission failure derives from an illicit switching of hypothetical contexts. In a normal context, it is suggested, one can get a warrant for "There are [only] zebras in that cage" by casual observation, and thereby—by the same casual observation—acquire a warrant for "There are not painted mules in that cage", by reasoning appealing to the additional premise that zebras are not painted mules. On the other hand, if the epistemic context is enlarged to accommodate a serious claim that the animals in question *may* be painted mules, then casual observation no longer suffices to confer a warrant on "There are zebras in that cage". So, depending on the context, we either have a transmissible warrant or no warrant. But in no case do we have a non-transmissible warrant.

It seems likely that a move of this general structure can be made in response to any case where it is suggested that there is a relatively unobvious form of question-begging going on. Rather than allow that warrant for the premises of a particular argument may presuppose some independent warrant, or entitlement, for its conclusion, it may always be insisted that the warrants for the premises divide into two classes, depending on the epistemic context. If the context is one in which the relevant conclusion is not explicitly considered, much less doubted, then to acquire an ordinary—"casual"—warrant for the premises will be, by inference, to acquire such a warrant for the conclusion too; but if the context is one in which the conclusion is explicitly contemplated, and a question about it needs to be addressed, then such "casual" grounds may not even add up to warrant for the premises.

It's not completely clear to me whether Sainsbury is indeed commending this move absolutely generally. The issues raised by the general epistemological contextualism to which, if I read him correctly, he here evinces an attraction are too big for discussion in this setting. I will make two brief observations. First, someone who is sceptical—as I am—about the general theoretical fruitfulness or efficacy of such contextualism is not of course committed to denying that what may count as possession of warrant for a particular claim may well depend on contextual variables of purpose, interest and need. Second, it cannot generally be the right thing to do to concede one's warrant for the premises of a given argument as soon as some serious question is raised about the status of the conclusion—after all, the argument may well be precisely what settles that question. So Sainsbury owes an account of the distinction between cases where it's unreasonable to see one's warrant for the premises of a given argument, together with its validity, as properly *stilling* an antecedent doubt about the conclusion, and cases where, rather, that antecedent doubt should cause a recognition that one's purported warrant for the premises is not good enough (as with the casual observation of the zebras). And it seems to me that will just turn out to be the very distinction to which I'm pointing: the former cases will be those where making out one's warrant for the premises will not involve defending an independent warrant/entitlement to the conclusion, and the latter cases will be those where it will—where, depending on how the attempt at an independent defence works out, one will indeed either wind up withdrawing the claim to have warrant for the premises (the only case Sainsbury envisages) or (the case he overlooks) making the independent warrant/entitlement out and thereby sustaining a *non-transmissible* warrant for the premises.

Even though sceptical about the idea of transmission failure, Sainsbury thinks there may still be some virtue in my second template as a tool for the diagnosis of what goes wrong with certain kinds of anti-sceptical argument. He points out—I think quite correctly—that the simple argument which persuaded Martin Davies (and me) that there was something wrong with Davies' initial formulation of his second limitative principle,³ viz.

A: I believe that water is wet; hence

B: someone is capable of believing something

can be made to conform to the second template if we add

C: no one is capable of believing anything,

envisaged as motivated by a philosophical eliminativism about the intentional. The general point which Sainsbury takes this to illustrate is that if the argument from A to B is directed against a considered *scepticism* about B, then that scepticism, if properly thought through, will involve making out how there is a type of situation—C—in which A fails but in which one's subjective state is indistinguishable from what it would be if A were true, and hence that one cannot just take it without further ado that A is true when one finds oneself in such a state. I think this is probably right, and would only question whether the presence of anything worth describing as "scepticism" is a necessary condition for the applicability of the template. One doesn't, after all, need to be a sceptic about the past in order to appreciate that there is something silly about offering apparent memories as evidence that there has indeed been an extended past. It is obvious that an entitlement to the past, so to speak, is a rational precondition of granting apparent memories the probative force for the specific claims about one's past which we routinely allow.

Sainsbury suggests that when one locates the point for which I was fumbling by my use of the A-B-C template in the general observation about the dialectics of anti-sceptical argument that he makes, one will see that Putnam's proof after all falls to the same consideration. Specifically: a sceptic about the conclusion of Putnam's proof ought not to allow that there is warrant for all of its premises. For

If the sceptical possibility about whether I am a brain-in-a-vat were realised, then the thought actually expressed by Putnam's premise (ii) ("in BIV-ese, 'brain-in-a-vat' does not refer to brains-in-a-vat") would not be available for use in any argument, since to suppose that it is warranted would be to suppose that the relevant scepticism is false, and so to beg the question.⁴

Thus, Sainsbury's thought is, just as a sceptic about my, or all my community's, history of previous encounters with water will not—in the presence of the relevant externalism—be willing to grant that I have warrant for the claim that I believe that water is wet (since the scepticism is about the availability to me of a conceptual resource needed by that claim), so a sceptic about whether I'm a brain-in-a-vat will not grant me warrant for the claim that BIV-ese does not contain the resources to refer to brains-in-vats, since—again—the scepticism is, in the presence of the relevant externalism, about a conceptual resource needed by that claim.

There is an analogy. But more care is needed, it seems to me, than Sainsbury takes about the content of the relevant forms of scepticism. If the Water Sceptic—perhaps an actual adversary, perhaps myself in a sceptical moment—insists that there is a well-taken doubt whether I have had the encounters with water requisite for grasp of the concept, he credits me—by dint of the very intelligibility to me of the allegedly well-grounded doubt—with a conceptual resource which, in the presence of the externalism, undercuts his question. And it's exactly the same with the Brain-in-a-Vat Sceptic: if the key question concerns with what right I take it that I'm not a brain-a-vat, the answer is, simply, that I cannot have a well-grounded doubt about *that* since the very grasp of the doubt demands conceptual resources which I couldn't have if the sceptical scenario obtained. Sainsbury's analogy thus obtains between two *foolish* forms of scepticism.

What are the smarter scepticisms? Well, the smarter Water Sceptic canvasses the possibility that all my seeming-tokens of a certain purported kind of thought involve illusion of content, and asks for a ground for discounting this possibility. That's a doubt I can grasp without presupposing that I possess the concept water. By contrast, the disanalogy with Putnam's argument which I was noting—and which Sainsbury seems to miss—is that there is no obvious way of formulating a "smarter" version of the doubt of the brain-in-a-vat sceptic. In the McKinsey case, the sceptical doubt concerns whether I have fallen prone to an illusion that I have a concept of a certain sort; in the case of brain-in-a-vat scepticism, the doubt *ought* to be whether my concepts are not those of a normal embodied human but, such is the supposed effect of the relevant externalism, the quite different set of concepts that would be possessed by a brain-in-a-vat. But *that* way of putting the doubt makes use of the concept—brain-in-a-vat—which is precisely one of those I wouldn't have if the doubt were realised. So it is *still* a foolish form of scepticism. The note on which my paper closed was that I did not—and still do not—see how to stop its being so (always provided that an appropriate externalism is true!).

Let me finally turn to Sainsbury's suggestion that there are problems with disquotation in both the McKinsey and Putnam arguments. He quite correctly observes that disquotation may not be a sound principle when applied within the scope of meaning-affecting hypotheses. For instance, if "pigs" referred to what "birds" currently refers to, then "'pigs fly' is true iff pigs fly" would still be true, but it would not be true that "pigs fly" is true iff pigs fly. However, the obvious obstacles to doing much with this observation are three:

that in the version of the Putnam argument I presented which explicitly relies on disquotation, it is unobvious that any application of it *does* take place within the scope of a meaning-affecting supposition;

that there is no obvious reliance on disquotation in the McKinsey argument;

that the play with disquotation in the Putnam argument seems in any case inessential and disappears when the argument is recast in the MC-form.

As far as the first of these is concerned, one might have expected that Sainsbury would argue that premise (ii) of the more longwinded version of Putnam's proof:

(ii) in BIV-ese, "brains-in-a-vat" does not refer to brains-in-a-vat,

tacitly incorporates some potentially meaning-affecting hypothesis, and is therefore potentially disabling of disquotation elsewhere in the argument. But Sainsbury gives no such argument, and it is indeed hard to see that he could have—for premise (ii) merely reports a consequence of the fact, purportedly enjoined by externalism and the details of example, that brains-in-a-vat will be debarred from referring to brains-in-a-vat. Rather, Sainsbury's argument seems to be that it is premise (i) that is question-begging—that my knowledge that my language is disquotational should, in this context, be regarded as entitling me to no more than that each instance of the disquotational scheme is a *true sentence*. I can know for instance that "'brain-in-a-vat' denotes brains-in-a-vat" is a true sentence, but I may not in turn presume—without begging the question—to know the result of disquotation on that sentence, viz. that "brain-in-a-vat" denotes brains-in-a-vat.

I can see no force in this. To fix ideas, consider a case of slow switching: the sceptical thought that I may have been transported a while ago to Twin Earth and may have been there long enough for my uses of "water" to have undergone a shift in reference. Whatever the facts, I know that "'water' refers to water" is a true sentence in the language I currently speak. Is there a problem nonetheless about disquotation? The sceptical hypothesis is certainly potentially meaning-affecting. But that seems to have no bearing on the soundness of disquotation: whatever the current semantics of my language, I am assured that a *use* of the expression, "water", will correctly specify what that expression stands for. Disquotation under a meaning-affecting hypothesis is

fine if the hypothesis is taken to be *actually* true. If my uses of “water” actually have shifted in reference over the last twenty years, that won’t affect the fact that by “water” I refer to: water. What is unacceptable is the counterfactual conditional: if my uses of “water” had shifted in reference over the last twenty years, then by “water” I would refer to: water.

This may all seem pretty academic since, as remarked, neither the McKinsey argument nor the Putnam argument in MC form seem to involve any play with disquotation. However, Sainsbury thinks these appearances deceptive. He envisages the externalist argument for the second premise in the McKinsey argument:

- (2) Anyone who believes that water is wet has had such and such encounters with water,

as proceeding via two lemmas: first that there is a class of concepts whose possession requires encounters with things that fall under them, and second (“a lemma whose a priori status may well be in doubt”⁵) that this applies to the concept labelled “water”. The result is accordingly

Anyone who possesses the “water”-concept has encountered things which fall under this concept.

And now to move from there to premise 2 *does* require a disquotation:

(D-W) Something falls under the “water”-concept iff it’s water.

Sainsbury’s next point is not that the disquotation principle (D-W) is suspect in this context—if I read him correctly—but rather that it’s not obvious that it is a priori. And of course, if some non a priori premise is tacitly at work in the McKinsey argument, then that suffices to defuse the paradox, which precisely consists in the apparent availability of a priori grounds for a historical conclusion. If the grounds for the premises are not wholly a priori, that appearance is just an illusion.

Sainsbury himself does not assert that (D-W) is not a priori, but only that there is an issue about the matter. For my part, I remain unpersuaded that homophonic disquotation into one’s own language is not something on which one may non-empirically rely. But even if I’m wrong about that, there is cause for thinking that Sainsbury’s diagnosis does not yet take us to the heart of the matter. For supposing it is indeed a priori that certain concepts, including the “water”-concept, are grasped only by those who have

encountered items in their extension, we still have a version of the McKinsey argument which runs like this:

- (1) I understand "water", and so possess the "water"-concept.
- (2) Anyone who possesses the "water"-concept has encountered things which fall under this concept.

Therefore

- (3) I have encountered things falling under the "water"-concept.

If understanding, no less than thinking, or believing, falls within the scope of privileged access—that is, if that one understands a given word, so possesses the concept expressed by it, something which one may normally be expected to know non-inferentially, then something essentially of the same form as the McKinsey argument remains even after what Sainsbury regards as the suspect play with disquotation is eliminated.

And there is in any case, finally, the question of how to apply Sainsbury's diagnosis to the Putnam argument in MC-form. (Here of course the dialectical situation is different: it would not suffice, in order to find a problem for Putnam's argument, to show that it involves a tacit reliance on disquotation and that the latter cannot be certified *a priori*. Even if it cannot be certified *a priori*, the proof can still be cogent provided disquotation is indeed known to be reliable *somehow*.) If we try to run the argument in a way which avoids any attempt to identify the concept expressed by "brain-in-a-vat", then presumably we get something like this:

- (1) I have the concept expressed by "brain-in-a-vat".
- (2) No instance of the concept expressed by "brain-in-a-vat" possesses that concept.

Therefore

- (3) I am not an instance of the concept expressed by "brain-in-a-vat".

Sainsbury's thought then seems to be that any advance from (3) to the conclusion that I am not a brain-in-a-vat involves a somehow question-begging use of disquotation—since it goes past the possibility, if I am a BIV, that the fact at which I should be arriving by disquotation is that I am not a *BIV-in-the-image* (am not in the extension of "brain-in-a-vat" as used in BIV-ese) though this is not how I "would or could put it".⁶ I can only repeat that I think this line of concern is completely wide of the mark. The sceptical worry was all along a worry formulated *using* the concept of a brain-in-a-vat, and there should be no squeamishness

about making use of that very same resource in the attempt to address it. The idea that any proper response must, on pain of question-begging, confine itself to metalinguistic formulations is totally unmotivated. And in fact it is clear that it would be impossible to make a good case for premise (2) as it stands without exploiting one's understanding that the concept expressed by "brain-in-a-vat" is indeed the concept, *brain-in-a-vat*. For the argument—of course Putnam was always rather inexplicit about the details—which would purport to show that (2) is true could not proceed just by general reflection on the kind of meaning possessed by "brain-in-a-vat" and the relevant semantic externalism. Rather, it must turn on the specific thought that *brains-in-a-vat*, in the situation of Putnam's example, would be debarred from forming the concept: *brain-in-a-vat*.

SUAREZ

Alfonso Suarez is unpersuaded that any deep distinction corresponds to the two templates for transmission failure—respectively typified by the soccer and McKinsey cases—that I distinguished. The mark of the first kind of case is information-dependence: the circumstance that one's evidence for the premises of the argument in question depends for its force on one's possession of collateral information including certain propositions which are themselves consequences of those premises. The mark of the second kind of case is that one's warrant for a premise of the argument is bestowed not by evidence—in the sense of independently statable information—but directly, by the operation of some appropriate sensory or intellectual faculty which acts as a receptor for facts of the appropriate kind. In such cases, the key thought as far as transmission failure is concerned is that such a faculty must be capable of certain kinds of misfire, so that in taking its operations to warrant a particular belief in particular circumstances, one presupposes a rational entitlement to the supposition that such a misfire is not at work, and hence that conditions which would suffice for such a misfire do not obtain. The thought is then that if some consequence, B, of the proposition which one takes oneself to have warrant for is such that (one knows that) were it to be false, that would suffice for a misfire of the appropriate faculty, one needs a prior rational entitlement to B in order to regard oneself as warranted in accepting the original premise. So the inference from that premise to B is not of the service of a first-time rational acceptance of the latter.

The two templates are thus designed for two quite different kinds of situation: respectively, the case where warrant for the premises of an inference is itself given by defeasible inference from independently statable evidence—and where Quinean considerations are thus available to condition the transition from that evidence to the premises—and the case where a warrant is rather the product of our presumed direct cognitive interaction, courtesy of appropriate faculties, with the facts depicted by the premises. With that difference properly emphasised, it is open to us to allow—as I am happy to do: I've said nothing to the contrary—that the two templates can be more closely assimilated than my discussion made explicit. Thus one would expect, as Suarez suggests, that it would generally be possible to represent instances of the first template under the rubric of the second. If in the soccer example, for instance, we take A as “A goal was just scored”, B as “A game of soccer is taking place”, and C as, say, “This is not a genuine game of soccer and these people are going through the motions for some other reason”, it would be no less plausible than in the zebra or deceptive lighting cases that each of the four conditions involved in the second template are met:

that A entails B;

that my warrant for A consists in my being in a state—witnessing the events on a field—which is subjectively indistinguishable from the state I'd be in if the relevant C were true;

C is incompatible with A; and that

C would be true if B were false.

And now the transmission questioning thought will just be that in order to regard myself as warranted in accepting A, I had better already be in a position to discount C and hence in a position to discount not-B. So I need a prior independent warrant to accept B—independent, that is, of the recognition of its entailment by A.

The reasoning associated with the second template is thus plausibly taken to be capable of explaining the transmission failures associated with instances of the first template. That is Suarez's first point of assimilation. His second suggestion is that if the warrants bestowed by the operation of faculties like perception and memory—in general, faculties presumed capable of direct reception of states of affairs of certain kinds—are still conditional on an independent entitlement, or warrant to take it that the

faculties in question are operating properly on the relevant occasion, then these warrants too should be regarded as information-dependent no less than those bestowed by defeasible evidence. For in both cases whether or not one is indeed warranted in accepting a certain claim is conditional on what else one is warranted (or entitled) in taking to be the case.

Again, I think this observation is fair as far as it goes—and indeed had not meant to imply anything to the contrary. More: the point that even warrants bestowed by the operation of directly receptive faculties are not unconditional but have this broadly information-dependent—perhaps better, *entitlement-dependent*—character is a point of considerable importance, and lies at the heart of my diagnosis of how inferences from non-inferentially warranted premises can nevertheless fail to transmit warrant. However—and this is my only serious disagreement with him—Suarez thinks that to acknowledge this point is to see that the McKinsey argument does after all bring out a clash between semantic externalism and our intuitive preconceptions about privileged access. For he seems to hold that to allow even this modest degree of “information dependence” in the warrants we have for claims about our own psychological states is somehow at odds with the intuitively immediate and non-inferential character of self-knowledge. So externalism, since it forces this kind of information dependence on our warrants for self-knowledge claims, is indeed ultimately in tension with privileged access as intuitively conceived.

This seems to me to be mistaken. When we, in our ordinary thinking, regard self-knowledge as, at least in basic cases, non-inferential, we have in mind no more than that it is in general neither possible nor necessary to produce corroborative evidence in favour of one's opinions concerning one's own intentional and sensational states. Even if some well-motivated form of semantic externalism opens up a previously unenvisaged way in which sincerely offered professions concerning one's own psychological states may be undermined, it will do nothing to compromise their non-inferential character, as just adumbrated, unless it somehow creates a need for *additional empirical work* before such claims can rightly be regarded as warranted. If that were the upshot, then a request for justification would not in general be inappropriate, and could be answered by describing the upshot of the necessary additional empirical work. That the effect of externalism is not to be viewed like that, however, was the principal point of the analogy—which Suarez for his part accepts—between the role of perceptual illusion in the defeat of warrants for ordinary percep-

tual claims and the role of content illusion in the defeat of warrants for claims about one's own psychological states. If it is right that these are not possibilities which it is necessary—or indeed generally possible—to exclude empirically before one can claim warrant for ordinary perceptual or psychological claims—if it's rather that we are a priori entitled in general to suppose that all is well in these respects—then the degree and kind of information-dependence which, as Suarez stresses, affects even the warrants bestowed by the operation of directly receptive faculties is quite consistent with the propriety of those aspects of our linguistic practice which express the ordinary idea that such warrants are non-inferential.

HALE

Bob Hale's careful and well-considered comments make a number of constructive points. I'll register and comment on four such before concluding with a point of disagreement.

First, Hale seems to be right to distinguish *strong* and *weak* closure principles, and to contrast both with transmission. What I called simply "closure" is Hale's weak closure: the principle that warrant is available for the premises of a valid argument only if it is available for the conclusion too. Unrestricted transmission says that, for any valid argument, a thinker may, by acquiring warrant for the premises and recognising the validity of the argument, thereby acquire warrant for the conclusion. So according to transmission, a sequence of steps of valid argument always corresponds to possible sequence of rational belief-formation, consequent on the acquisition of warrant for the premises and recognition of the validity of the argument step by step. *Strong* closure, by contrast, says, roughly, that whenever there is warrant for the premises of a valid argument, the *very same* warranting considerations, augmented only by recognition of the validity of the argument, will constitute warrant for the conclusion. Hale's well-taken observation is that strong closure does not suffice for transmission, since it may still be, even in a case where strong closure holds, that the considerations which warrant the premises of the argument do so only because they warrant the conclusion *independently*—for instance, if the warrant for the premises is an information-dependent warrant, with the relevant conclusion included in the needed pool of information, while at the same time warranting (perhaps in a way depending on different information) the conclusion itself.

A simple example of such a scenario is provided by a modification of the soccer example whereby, rather than relying on the testimony of a bystander that a game of soccer is in progress, my background information is such that I am in a position to treat the very evidence of my eyes as evidence for that claim. With that information in place, the witnessed scene can once again provide a warrant for the proposition that a goal was just scored. But as before, it is only because I have, in the example as it now is, observational evidence that a game is taking place at all that the witnessed scene provides evidence of the goal: the warrant for both premise and conclusion is the evidence of my senses, but there is still a failure of transmission—I do not learn of the conclusion by acquiring a warrant for the premise and carrying out the inference, but rather have a warrant for the premise in the first place only because the conclusion is already in place, warranted by the very same evidence.

Second, Hale is also quite right to interpret my conception of information-dependence as implicating a broadly internalist conception of warrant. That is to say, the idea of a warrant's being information-dependent is meant to import not a relativity to what happens to be, merely, *true* in the context in which a thinker is working but a relativity to what she independently has evidence for or is otherwise entitled to believe. This is, in effect, the notion that is at work in Quine's classic repudiation of the traditional category of *synthetic* statements, when—as he argues—the confirmatory propensities of experience vis-à-vis any particular statement are dependent on background *theory*. It is actually crucial for the connection between information-dependence and subtle question-begging that the former be taken in this broadly internalist way. For if *e* was an *I*-dependent warrant for *P* only in the sense that it was a necessary condition for *e*'s warranting *P* that *I* be *true*, then, even in the case where *P* entails *I*, there is no evident reason why someone who possessed what was in fact a warrant for *P*—viz. *e*—could not come to a warranted belief that *I* just by carrying out the inference. (Of course it might still be that in order to *recognise* that they were warranted in believing *P* in the first place, they would need to recognise that *I* obtained; but that would be a condition on having a *second-order* warrant for *P*—warrant for thinking that *P* was warranted by *e*—and would pose no barrier to the transmission of the first-order warrant. It is of course a characteristic of externalists to be insouciant about the absence of second-order warrants!)

Third, Hale makes the following, as it seems to me correct, observation about the relations between strong closure—as so far characterised—and transmission. The observation assumes, as seems plausible, that information-dependent warrant obeys this (I'll call it *Hale's*) principle: that when a thinker has a warrant, *e*, for *P* depending on a body of background information *I*, then *I* will not suffice on its own, irrespective of her possession of *e*, to warrant *P*. It follows immediately that the warrants the thinker has for each proposition *Q* in *I* cannot themselves be dependent on the same body of information, *I* (for if they were, that information would, trivially, warrant each such *Q* in its own right, which would be a counterexample to the stated principle). So we have the following consequence of Hale's principle: if strong closure is formulated as requiring that whenever there is an *I*-dependent warrant for the premises of an entailment, that same evidence, depending on the *same* information, should constitute a warrant for the conclusion, then strong closure is going to fail in all the cases where some components of *I* are consequences of *P*—exactly the cases where, as I argued, information-dependence leads to failures of transmission.

The immediate effect is to highlight the need for a further distinction, this time among *strong* closure principles. For the strong closure principle just noted must therefore fail even in the refashioned soccer case, when the same body of perceptual evidence warrants both the claim that a goal has been scored and the claim that a game of soccer is in progress. And indeed it does. For while that is—so I suggested—an example where the evidence for premise and conclusion is the same, it's plausible that that evidence in both cases confers an information-dependent warrant, and that the relevant bodies of background information are *not* the same. In particular, the body of information subject to which the evidence confirms "a goal has just been scored" must include the proposition that a game of football is in progress; whereas that proposition is not, of course, part of the needed background information if the evidence in question is to confirm *it*. Thus: strong closure as Hale understands it demands that any information-dependent warrant for the premises of a valid argument amount to a warrant, *relative to the same information*, for the conclusion of that argument; whereas the weaker—though still strong—principle requires only that a given body of evidence constitutes a warrant for both premise and conclusion, allowing that the information on which it depends, as it discharges those respective roles, may differ.

The question all this raises, of course, is whether Hale's version of strong closure in effect captures everything essential to the transmissibility of information-dependent warrants: whether, that is, any counterexample to transmission of information-dependent warrant is a counterexample to strong closure in Hale's sense and vice versa. If so, that might, to be sure, be merely an extensionally adequate characterisation of transmission—it might give us no additional insight into its nature. But the equivalence would still represent a useful control. I leave the issue open here.

Fourth, the issue concerning the relationship between transmission and strong closure principles looms large again in connection with the McKinsey argument. If I am right, *pace* Suarez, that some kinds of warrants—those afforded by the operations of faculties, like perception, memory, introspection or logical intuition which are thought of as directly, non-inferentially receptive to certain kinds of states of affairs—are not usefully regarded as information-dependent, then the distinction just made between two kinds of strong closure principle will not arise. Rather, we will be left with an undifferentiated notion according to which strong closure is satisfied just when in order for the thinker to have a warrant for the conclusion, nothing beyond recognition of the validity of the inference needs to be added to the deliverances of the relevant faculties that constitute the warrant for the premises of the argument. There is, indeed, some awkwardness about expressing the matter since we now have nothing to correspond to either *e* or *I* in the information-dependent setting. So the notion of having the “same” warrant limps. Just for that reason, it's hard to know now how to formulate even an initial distinction between strong closure and transmission. What is clear, though, is that the McKinsey argument poses a paradox on the assumption that it obeys either principle. One naturally thinks of transmission as being the crucial thing, since the paradox turns on the idea that one can get a warrant for the historical conclusion of the argument just by the broadly reflective means involved in ratification of its premises. But strong closure, as formulated a moment ago, would still allow that the operation of these reflective faculties—together with reflection on the entailment—can generate a warrant for the historical conclusion. I believe that the response which I suggested in my paper lends itself equally to the explanation of why strong closure, intuitively understood, fails for the McKinsey argument. If one does indeed possess a non-empirical warrant for both premises in the McKinsey argument, it is courtesy of an a priori entitlement to discount the possibility of illu-

sions of content, and hence to discount any scenario that would generate such an illusion in the relevant case. Even if one so uses “warrant” as to embrace this kind of entitlement, as well as more specific cognitive achievements, it remains that this warrant is not a product of the operation of the faculties from which one’s warrants for the psychological and externalist premises eventuate. So strong closure fails.

In conclusion, finally, to the advertised point of disagreement. Hale is inclined to accept my diagnosis that Putnam’s proof is not afflicted by the failure of transmission, or closure, involved in the McKinsey argument. But of course it does not follow from its not having that particular problem that the proof is, as I suggested, cogent. Hale thinks it has another problem: specifically, that it is still a question-begging argument, even though not in any of the ways in which I was concerned.

His reservation is, indeed, one that I’ve discussed elsewhere, and here I must be content with the briefest indication of why I do not share it. The worry is, roughly, that it goes with the sceptical fantasy that we are trying to displace—the idea that I might actually be a brain-in-a-vat in a universe consisting in nothing but brains-in-vats, attended by automatic machinery, etc., just as Putnam described—that this whole dialectic, all the thoughts that we are having now, are nothing but thoughts of brains-in-vats. So a brain-in-a-vat can run Putnam’s proof—in BIV-ese, which is the language we are actually thinking in if we are brains-in-vats—and validly arrive at a conclusion which, whatever it is, cannot be the conclusion that it is not a brain-in-a-vat, since that is just what it is. In short, in order to be sure that Putnam’s proof delivers the intended conclusion, I need to know—in advance—that the “I” who performs the argument is a normal embodied human and not a brain-in-a-vat.

I think there *is* a legitimate worry of this general kind in cases where the sceptical fantasy involves, for externalism-driven reasons, unmonitored switching of meanings. Suppose for example I am disturbed by the thought that agents may secretly—in my sleep—be shuttling me back and forth between Earth and Twin Earth, allowing me to stay long enough in each place to bring about reference-shift in the terms “Water”, “Earth” and “Twin Earth”. Clearly I cannot reassure myself by running this argument:

(i) In my language, “water” denotes water

(ii) In the language spoken on Twin Earth, “water” denotes XYZ

Hence

- (iii) My language is not spoken on Twin Earth
 Hence
 (iv) I am not on Twin Earth.

Here it would be absolutely right to say just what Hale does say: that the content of this valid argument will vary as a function of my situation when I run it; hence that it is powerless to assure me that my sceptical worry is ungrounded. But it is absolutely crucial to this objection that the sceptical fantasy is one of recurrent *meaning change*. Putnam's scenario, by contrast, is one in which the reference of the crucial expressions—in particular "brain-in-a-vat"—is supposed to be *constant*. The world has *always* consisted of brains-in-a-vat attended by automatic machinery and nothing else. And what we have been challenged to do is to produce a reason for thinking that when the scenario is put to us—in *our language* (what else?)—it is not being put to creatures of whom it is actually true. We are supposed to *understand* the fantasy, and the words being used to express it are not—according to the fantasy—up for variation in meaning on different occasions of contemplation. *Brain-in-a-vat* is a concept which we are supposed to have—we need it to understand the fantasy—and the question is: is it consistent with everything that we are allowed to know or understand, as part of our understanding of the challenge being put to us, to suppose that we fall under this concept? And the answer is that it is not—at least not if the example is properly constructed and some relevant form of externalism is sound. Indeed, the example is precisely to be so constructed that creatures who fell under this concept *couldn't themselves possess the concept*. We, who grasp the challenge, *do* possess the concept. So we don't fall under it.

In the paper, I added an ironic proviso: "*sotto voce*: whichever concept that is". Clearly I didn't make the irony explicit enough. Certainly, there are two concepts in the offing—the concept of a brain-in-a-vat, possessed by normal human beings who understand Putnam's scenario; and the concept—if there could be such a thing—expressed by the BIV-ese expression, "brains-in-a-vat". But in order for Hale's misgivings to kick in, it has to be uncertain which concept of those two we token by "brain-in-a-vat". And it isn't in the least uncertain. The concept we possess is the one you need to understand the sceptical challenge; and that is the concept which it then turns out creatures who fell under it—precisely: brains-in-vats—couldn't possess. We do possess it. So our concept cannot be the concept expressed by "brain-in-a-vat" in BIV-ese. QED.⁷

NOTES

1. This volume, p. 192
2. This volume, pp. 193–194.
3. That warrant cannot be transmitted across a known entailment if the truth of the consequent is a precondition of the subject's capacity to believe the antecedent. See above, this volume p. 147.
4. This volume p. 195.
5. This volume p. 198.
6. This volume, p. 199.
7. Thanks to Sven Rosenkranz.