

Internalism and the snapshot conception of phenomenal experience: a reply to Fisher

Gary Bartlett, Central Washington University

1. Introduction

Over the past forty years there have arisen two main lines of resistance to the view that a subject's mental state supervenes on her internal physical state. The seeds of the first line were sown by Hilary Putnam (1975) and cultivated in particular by Tyler Burge (1979; 1986). Putnam, Burge and their sympathizers argued that since the *content* of a mental state had to constitutively depend on the subject's environment, and so could not supervene on her internal states, and since content individuates a subject's mental states in the first place, then the mental states themselves could not supervene on the subject's internal states.

The second line of resistance is of more recent vintage. In the late 1990s some philosophers, led by Andy Clark and David Chalmers (1998) and Susan Hurley (1998), began to argue that the mind extended beyond the body in a more literal fashion. To put it in Hurley's terms, the view held that the vehicle of content can sometimes be located partly beyond the body. Clark and Chalmers called the view 'active externalism' to emphasize that certain kinds of active causal engagement between internal and external factors could lead to a subject's beliefs, in particular, residing partly (or even fully) outside her body.

I offer this very brief potted history as a lead-in to a discussion of what seems to me to be an attempt at a novel, third line of resistance to internalism: that of Justin Fisher (2007). Fisher's reasoning does not fall neatly into either of the two existing externalist camps. It is not content externalist, as it does not focus on mental content. Nor is it vehicle externalist, as it does not claim that mental states extend beyond the skin. Yet it shares the former's strategy of denying that physical indistinguishability entails mental indistinguishability, and the latter's emphasis on the active dependence of mind on environment. Fisher's moral is that "the normal functioning of *all* cognitive systems deeply depends upon their getting appropriate support (or at least

appropriate non-interference) from their surroundings” (p. 324). This means that if one knows only the current internal physical features of a system, it is impossible to determine its mental features. Fisher claims to show that internalism is false of *all* mental features – not just cognitive states such as beliefs, which are the usual targets for both content and vehicle externalists, but also “phenomenal experiences, rationality, moral character, emotion-types, and propositional-attitude-types” (p. 321). His article’s bold title is ‘Why nothing mental is just in the head’.

In what follows I shall work with the most common definition of internalism: that mental states supervene on brain states. This is the definition that Fisher assumes.¹ My particular concern is with internalism about phenomenal experiences – that is, *phenomenal internalism*. I shall contend that Fisher fails to refute phenomenal internalism. That is, I shall contend that he fails to show that phenomenal experiences are not just in the head.

2. Fisher’s argument

Fisher intends his argument to refute internalism across the board, concerning all kinds of mental features. I shall first describe his argument, and then offer a more formal reconstruction of it for purposes of clarity.

Fisher deploys a case concerning ‘Pulse World’, which is a planet orbiting a pulsar. Pulsars are stars that emit radiation bursts at regular and rapid intervals. Pulse World’s pulsar emits 100 pulses per second. An Earthling who was subjected to the pulses, Fisher has us suppose, would immediately begin to think and act in conspicuously irrational ways, for the pulses would dramatically interfere with her brain activity. The denizens of Pulse World, however, are well-adapted to the pulses; indeed, they depend on them. Were the pulses to cease, Pulselings’ brain activity would then result in thoughts and actions that would be just as conspicuously irrational as we Earthlings would display if we were subjected to the pulses.

¹ In my 2008 I argued that this definition needs revision to allow for the possibility of the causal basis of the mind extending beyond the body, but nothing will turn on that matter here.

Fisher compares a Pulseling, Paula, and an Earthling, Edna, at a given moment. For convenience I shall call that moment t_1 . t_1 is between pulses on Pulse World. At t_1 Paula is driving on a freeway, and is surprised that her exit has already arrived. Meanwhile, Edna is playing the saxophone, and is anxious that she will play a bad note. Their personal histories are quite distinct, as are their background beliefs and so on. Their physiologies are also very distinct, for Paula is not even humanoid – Fisher mentions that she has tentacles. In general, then, all the evidence strongly suggests that Edna and Paula are radically mentally distinct.

However, Fisher stipulates that “at this moment [that is, t_1], Edna and Paula happen to be completely indistinguishable with respect to what’s in their heads” (pp. 321-322). The pulses make this possible. While Edna’s brain activity *after* t_1 would not be suitable to control Paula’s body after t_1 – due to the distinctness of their bodies, environments, and behavior – the incoming pulses will nevertheless ensure that *Paula*’s brain activity evolves in a way that is suitable to exert that control. Thanks in part to the pulses, Paula will continue to drive her vehicle in an appropriate manner, rather than trying to play an invisible saxophone! The fact that Edna’s and Paula’s brains are “momentarily indistinguishable” (p. 322) at t_1 , then, is a mere accident. If we mapped their brain states in a multi-dimensional space, they would be on distinct trajectories which merely happen to cross at t_1 . Given how different Edna and Paula are in all other respects, Fisher urges, this momentary neural indistinguishability surely does not change the fact that they are radically mentally distinct, even at t_1 itself. Yet since they do have physically indistinguishable brain states at t_1 , then it seems that internalism must insist that they have indistinguishable mental states at t_1 – which is surely not the case. So internalism is false.

I reconstruct Fisher’s argument as follows:

- (1) Edna and Paula have physically indistinguishable brains at t_1 .
- (2) Internalism concerning a type of mental feature M entails that if two individuals have physically indistinguishable brains at a given moment, they must be indistinguishable with respect to M at that moment.

Therefore,

- (3) Internalism entails that with respect to any type of mental feature M, Edna and Paula must be indistinguishable with respect to M at t_1 . [from (1) & (2)]

But,

- (4) Edna's and Paula's behaviors, histories and situations at and around t_1 are radically distinct; and further, their brains are distinct both before and after t_1 .
- (5) If two individuals are radically distinct in their behaviors, histories and situations at and around a given moment, and if their brains are distinct both before and after that moment, then the two individuals are distinct in all of their mental states at that moment.

Therefore,

- (6) Edna and Paula are distinct in their content, justification, phenomenal experiences, rationality, moral character, emotion-types, and propositional-attitude-types at t_1 . [from (4) & (5)]

Therefore,

- (7) Internalism is false of content, justification, phenomenal experiences, rationality, moral character, emotion-types, and propositional-attitude-types. [from (3) & (6)]

The general thrust is that the physical indistinguishability of Edna's and Paula's brains at t_1 has no mental significance. It is a mere coincidence – literally! The momentary *co-incident* of their neural events occurs in the context of distinct neural trajectories, in accord with their very distinct mental and behavioral trajectories. Edna and Paula are thus mentally distinct even at t_1 , contrary to what internalism says.

3. Initial discussion of the argument

I do not dispute the possibility of the case that Fisher describes. I agree that Edna and Paula might momentarily occupy indistinguishable brain states even though they are engaging in the very different activities of playing the saxophone versus driving a car. For one thing, given their

very different bodies, the same momentary brain state will relate to distinct bodily states in each of them. But even if they had similar bodies, a momentary ‘snapshot’ of their brains might still show indistinguishable brain states. Behavior is a product not just of an organism’s brain, but of a dynamic interplay between its brain, its body and its environment. A single global brain state could feature momentarily in the course of two distinct behaviors.

I do not, then, reject the set-up of the case. What I reject is the claim that Edna and Paula are entirely mentally distinct at t_1 : in my formulation above, (6). While I agree that they are mentally distinct at t_1 in very many ways, I deny that they are distinct in their phenomenal experiences at t_1 . I deny this because I deny the crucial premise of the argument: premise (5).

Surprisingly, Fisher gives no defense of (5). He says only that his “descriptions of Edna’s and Paula’s respective behaviors, histories and situations make it clear that their mental lives must be radically different” (p. 321) in respect of all the mental features he mentions.

Now in point of fact, Fisher’s descriptions actually *stipulate* that Edna’s and Paula’s mental lives are radically different at t_1 . He directly describes only their mental features. Descriptions of their behaviors, histories and situations are found only within intensional contexts. For example, we are told that “Paula occurrently believes that she is driving her car”, while “Edna occurrently believes that she is playing saxophone” (p. 321). I shall assume, however, that Fisher did not intend to stipulate that Edna and Paula are mentally distinct, for his argument would then beg the question. His intention, then, must have been to stipulate only their *non*-mental features – hence premise (4) in my formulation of his argument – with the idea being that in light of the great differences in those features, it would be obvious that Edna and Paula must differ just as greatly in their mental features – hence (6).

That inference, however, must be mediated by (5). Since Fisher does not defend (5), he must find it to be in some way intuitive or self-evident. I do not find (5) to be intuitive or self-evident, and I am not convinced that it is true. Specifically, I am not convinced that two individuals who differ at a given moment in all of the cited non-mental ways must therefore be distinct in their phenomenal experiences at that moment. The proposition that the individuals *do* differ in their

experiences at that moment follows only with the addition of a controversial further principle. That principle is implicit in Fisher's argument. I shall now make it explicit.

3. Introducing the snapshot conception of phenomenal experience

Phenomenal experiences are my concern in this paper, so I shall now focus on them. Premise (5) assumes that at the moment in question the two individuals *have* mental states, and thus that they have phenomenal experiences. It therefore assumes the following principle:

The snapshot conception of phenomenal experience: Given a subject S who has a phenomenal experience ψ which lasts from t_0 - t_n , and given an instantaneous moment t_1 falling between t_0 and t_n , S must have some phenomenal experience ϕ precisely at t_1 (where ϕ is either identical to ψ or a part of ψ).²

This principle does not entail that all of the momentary experiences during t_0 - t_n are qualitatively indistinguishable. It says only that some experience or other is instantiated at each moment. So even if ψ varies qualitatively over time, it is still, as we might say, *phenomenally continuous*. A 'snapshot', metaphorically speaking, of S's mental features at any moment of t_0 - t_n would show the instantiation of some experience or other.³

² I have formulated the snapshot conception to be neutral as to whether an extended mental state is *composed of* a series of momentary mental states. There are two interpretations of why, according to the conception, there must be some mental state ϕ at t_1 . The first, which we could call 'perdurantist', says that since ψ is temporally extended, it must have temporal parts, and so ϕ must be one of them. This interpretation entails that all extended mental states are just a series of momentary mental states. But the other interpretation, which we might call 'endurantist', says that ϕ is *identical* to ψ : it is just ψ itself as it existed at t_1 . That is, ψ (or ϕ) is a single entity, having no temporal parts, but persisting through time as a full-fledged mental state at all moments. The difference between these two interpretations will not matter for my purposes.

³ I borrow the name 'snapshot conception' from Alva Noë (see especially Noë 2004). On Noë's enactivist view of perception, perceptual experience is an active *process* of engagement with the environment. He opposes something that he calls 'the snapshot conception' of experience. I do not claim that my formulation is exactly what Noë has in mind – in particular, he usually focuses on *perceptual* experience – but it is at least fairly close. It seems to be

[Footnote continues on following page]

For example, suppose that Ted, having spent too long on a merry-go-round, feels dizzy for two seconds. The snapshot conception says that this phenomenal experience of Ted's consists in his having some phenomenal experience or other at every instant within that two-second period. His experience in that period may or may not vary qualitatively, but it is phenomenally continuous: it contains no breaks in phenomenal character.

The 'snapshot' metaphor can become more literal if we think of functional brain imaging technology. The snapshot conception suggests that if it became possible to take a thoroughly detailed scan of a subject's brain at any given moment, we could tell exactly what experiences the subject was having at that very moment. Indeed, some cognitive neuroscientists – and perhaps some philosophers of a strongly internalist bent – may suppose that such a thing will indeed be possible if we can overcome the present limitations of the technology.⁴

It is easy to assume that such a thing must indeed be possible, technology willing. However, matters are not so simple. If the snapshot conception is false, we will no more be able to read off a subject's momentary experiences from a snapshot of his brain than we are able to read off a caterpillar's metamorphosis into a butterfly from a snapshot of a chrysalis, or a tsunami from a snapshot of a wave topping a seawall. For if the snapshot conception is false, then experiences

central to his view that experiences are dynamic and therefore temporally extended. Hence, he has claimed, there is no such thing as a momentary experience (see Noë 2006, p. 430). The snapshot conception that he opposes, then, is the view that there *is* such a thing as a momentary experience. This is what the snapshot conception of phenomenal experience, as I have formulated it, says is the case with phenomenal experiences in general.

However, Noë often claims that internalism entails, or at least strongly suggests, the snapshot conception. In Noë and Thompson (2004) he says that “the entire internalist, Cartesian tradition... rests on the assimilation of perceptual experience to a snapshot-like, phenomenal episode in the mind-brain” (p. 23). I find this claim wholly without merit, and have argued against it elsewhere (McLaughlin & Bartlett, 2004, p. 66). Therefore, while I am inclined to reject the snapshot conception of phenomenal experience, I do not think that this should incline me to reject internalism! My express purpose in this paper is, of course, to *defend* phenomenal internalism.

⁴ In particular, at the moment there is a trade-off between temporal and spatial resolution. While functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) offer good spatial resolution but poor temporal resolution, electroencephalography (EEG) and magnetoencephalography (MEG) have the reverse profile.

are like the metamorphosis and the tsunami in being *processes*, which are instantiated only *over* a period of time, and which therefore cannot be captured in a snapshot. Of course, someone who is familiar with the butterfly life-cycle may infer the rest of the metamorphic process just from an image of the chrysalis. But it would be a part-whole confusion to then conclude, of some temporal part of the metamorphic process, that it itself must also be a metamorphosis. Not all (and perhaps not any) temporal parts of a metamorphosis are themselves metamorphoses. Not all temporal parts of a tsunami are themselves tsunamis. And, if the snapshot conception is false, not all temporal parts of a phenomenal experience are themselves phenomenal experiences – even if a suitably knowledgeable neuroscientist might conceivably be able to infer the entirety of an experience just from a brain image of one moment of it. The experience may in fact arise only from the entire process, and hence may not be found at any single moment during that process.

Consider again the case of Ted, who experiences dizziness for two seconds. This does not entail that any temporal *part* of that experience must itself be an experience. Moreover, the snapshot conception concerns not just parts of experiences but specifically *instantaneous* parts of experiences. Even if some parts of experiences are indeed also experiences (e.g., perhaps the first half of Ted's experience is also an experience) it is a very much stronger proposition to say that a *moment* of an experience is also an experience – and indeed, that this is so of *all* moments of all experiences. Why think that there must be, at every instant of those two seconds, something which is an experience for Ted – whether of dizziness or anything else?

The snapshot conception gets its intuitive pull from introspection, I suppose: it just *seems* to us that our experiences are phenomenally continuous. But I see no reason to think that our introspective capabilities are sufficiently acute to justify that judgment, and I believe that there are compelling reasons to think that experiences are processes. So I am myself inclined to think that the snapshot conception is false. However, I shall leave that matter for another occasion, for I do not need to show here that the conception is false. I shall argue that Fisher's argument falls short *whether or not* the snapshot conception of phenomenal experience is false.

4. Why Fisher's argument fails

Let us return, then, to Edna and Paula. As a reminder, here again is the crucial premise (5):

- (5) If two individuals are radically distinct in their behaviors, histories and situations at and around a given moment, and if their brains are distinct both before and after that moment, then the two individuals are distinct in all of their mental states at that moment.

In this section I argue as follows. On the one hand, if the snapshot conception of phenomenal experience is false then (5) is false. On the other hand, if the snapshot conception is true then it becomes an open question whether (5) is true – so Fisher's argument is incomplete until he provides a defense for (5). I therefore aim to pose a dilemma for Fisher: whether or not the snapshot conception is true, his argument fails to disprove phenomenal internalism.

Here is the first horn of the dilemma. Fisher's claim that Edna and Paula are *distinct* in their experiences precisely at t_1 presupposes that they *have* experiences precisely at t_1 . Now this presupposition rests on the snapshot conception's assertion that since t_1 falls between the inception and termination of certain token experiences in Edna and Paula, then they must be having *some* experience or other at t_1 itself. If the snapshot conception is false, then premise (5) is false, simply because the two individuals may not have any experience at all at the moment in question. So if we reject the snapshot conception, we may deny that Edna and Paula have *any* phenomenal experiences at t_1 , and thus deny that they are *distinct* in their experiences at t_1 .⁵

Recall Ted and his episode of dizziness. Just as in that case, although Paula is, at t_1 , in the midst of an experience of surprise, this does not entail that there is, at t_1 itself, something which is an experience of surprise for Paula. *Mutatis mutandis* for Edna and her experience of anxiety.

⁵ Note carefully what premise (2) says: 'Internalism concerning a type of mental feature M entails that if two individuals have physically indistinguishable brains at a given moment, they must *be indistinguishable with respect to M* at that moment' – not that they must *have indistinguishable states* of type M at that moment. One way for two individuals to be indistinguishable with respect to M is, of course, by *neither individual having* any state of type M.

They may both be, at t_1 itself, phenomenally blank – and thus, contrary to (6), not distinct in their phenomenal experiences at t_1 .

We now turn to the second horn of the dilemma. What if the snapshot conception is true? It might then seem that Fisher's inference goes through: for surely, so long as Edna and Paula do in fact have some experience or other at t_1 , their experiences must be distinct. However, I think that this is not necessarily the case. The snapshot conception itself opens up the possibility that Edna and Paula undergo not just a momentary co-incidence in their neural states (despite their neural differences before and after t_1), but also a momentary co-incidence in their *phenomenal* states (despite their *phenomenal* differences before and after t_1). We may find this a strange idea, but the strangeness comes from the snapshot conception, which we are now embracing.

Fisher effectively suggests that given Edna's and Paula's enormous differences in behavior, situation, and history, and also the differences in their brain states before and after t_1 , their momentary neural indistinguishability at t_1 is just a neurophysiological co-incidence, which can have no significance for their mental lives. Hence his conclusion that they must have distinct experiences at t_1 . However, an internalist could respond that Edna's and Paula's momentary neural indistinguishability entails a momentary *phenomenal* indistinguishability, and that just as the former is a mere neurophysiological co-incidence with no significance for their mental lives at the time, the latter is a mere *phenomenal* co-incidence with no significance for their *ongoing* mental lives. Perhaps, that is, (5) is still false, and so Edna and Paula do indeed have exactly the same phenomenal experience precisely at t_1 .

We need not suppose either that Edna would experience a 'flash' of surprise in the midst of her anxiety or that Paula would experience a 'flash' of anxiety in the midst of her surprise. It is the thought of something like this, I believe, that causes us initially to recoil from the idea that they might have the same experience at t_1 . Closer consideration, however, reveals two points. Firstly, even if it *was* the case that Edna had a flash of surprise or Paula a flash of anxiety at t_1 , it is plausible that they would not even notice it; or that if they did (in some sense) notice it, they would simply forget it within the next few milliseconds, and thus that it would have no ongoing

effect on their conscious mental lives. But secondly, it is in any case far more likely that the momentary experience would be *neither anxiety nor surprise*. Rather, it would likely be some more basic experience which is a simple phenomenal constituent of both of the complex experiences of anxiety and surprise (or at least, which is a constituent of these two tokens of those experiences). And again, it would be an experience that the subject would not notice at all in the normal course of things – any more than one notices individual frames in a movie.

You might still find this story a bit hard to swallow. I would not blame you. I myself find it a bit hard to swallow. But that is because, as I noted earlier, I am inclined to think that the snapshot conception is false. It is, I think, a pretty implausible principle, and we should not be surprised if implausible principles have implausible consequences.

If Fisher thinks that not even the snapshot conception of phenomenal experience makes possible the momentary phenomenal co-incidence of Edna and Paula at t_1 , he owes us an argument for why that co-incidence cannot happen. In effect, he owes us an argument for (5).

Of course, it is not *just* the snapshot conception that entails the odd-sounding result that Edna and Paula have the very same experience at t_1 , but rather the conjunction of the snapshot conception and phenomenal internalism. So Fisher, finding the result not just odd-sounding but untenable, may insist that if we cannot reject the snapshot conception then we must reject phenomenal internalism. However, he cannot do this while simply taking the snapshot conception for granted. For as I pointed out in §3, it is not obvious that the snapshot conception is true; and as I pointed out earlier in this section, if the snapshot conception is false then Fisher's argument collapses. Thus to get us to reject phenomenal internalism rather than the snapshot conception, Fisher would need to convince us that the latter is more plausible than the former.

Let me sum up this section. For (5) to have any chance of being true, Fisher must assume the snapshot conception of phenomenal experience. If he does not, then there is no reason why Edna and Paula would have *any* experiences at t_1 itself, let alone radically distinct ones – so internalism will be correct that they will be indistinguishable with respect to their experiences at t_1 . On the other hand, if Fisher assumes the snapshot conception, he owes us a defense of (5), for

the conception itself permits the internalist counter-claim that the momentary neural co-incident entails an equally momentary phenomenal co-incident. This latter position is counter-intuitive, but that is no argument against it, since it may be that we are just not used to taking the snapshot conception seriously enough to see its full consequences. Either way, Fisher's argument fails to defeat phenomenal internalism.

However, there is one more concern to address.

5. Does Fisher's argument really depend on the snapshot conception?

You might wonder whether Fisher really needs the snapshot conception of phenomenal experience. Might not his stipulation that the neural co-incident lasts only a *moment* be strictly incidental to his argument, and could it not be abandoned?

To test this idea, let us suppose that the co-incident lasts for a period long enough to allow the determination of an experience. I shall show that even with this revision, the argument still fails. The ensuing discussion also affords us a broader understanding of the crucial but tacit role that the snapshot conception plays in making Fisher's original argument appear cogent.

We do not know how long a period is needed for a determinate experience to occur. This is an empirical question. But we need not stipulate any particular period. Let λ be the minimum length required.

Imagine a revised thought experiment in which the pulses come less frequently (e.g., one per second), thus allowing Edna's and Paula's brains to be indistinguishable for a longer period. If Fisher's argument does *not*, in fact, depend on the snapshot conception, then the resulting argument will be sound. Here is the revised argument:

(1') Edna and Paula have physically indistinguishable brains for a period t_1 - t_2 , which is of at least length λ .

(2') Internalism concerning a type of mental feature M entails that if two individuals have physically indistinguishable brains for a period of at least length λ , they must be indistinguishable with respect to M during that period.

Therefore,

(3') Internalism entails that with respect to any type of mental feature M, Edna and Paula must be indistinguishable with respect to M from t_1 - t_2 . [from (1') & (2')]

But,

(4') Edna's and Paula's behaviors, histories and situations are radically distinct at and around the period t_1 - t_2 ; and further, their brains are distinct both before t_1 and after t_2 .

(5') If two individuals are radically distinct in their behaviors, histories and situations at and around a given period, and if their brains are distinct both before and after that period, then the two individuals are distinct in all of their mental states during that period.

Therefore,

(6') Edna and Paula are distinct in their content, justification, phenomenal experiences, rationality, moral character, emotion-types, and propositional-attitude-types from t_1 - t_2 . [from (4') & (5')]

Therefore,

(7) Internalism is false of content, justification, phenomenal experiences, rationality, moral character, emotion-types, and propositional-attitude-types. [from (3') & (6')]

This revised argument is not sound. Given the constraint imposed by (1'), (4') cannot be true in way that will confer plausibility on (5'). Let me explain.

First, notice that merely slowing the pulses will not extend the neural co-incidence. For as Fisher himself emphasizes, Edna's and Paula's brains also receive input from their bodies, activities, and environments. These 'external features', as I shall call them, are quite distinct between Edna and Paula, and will quickly cause their brains to diverge. Arranging for the two brains to remain indistinguishable for more than a moment will require serious jury-rigging.

(This is already a hint that the argument has a more than incidental reliance on the snapshot conception.)

One way to achieve a longer neural co-incidence would be to make Edna and Paula much more similar in their external features (their bodies, activities, and environments). But those features would have to be *very* similar, if not indistinguishable, if they are not to cause the evolving brain states to diverge. And such great similarity in the external features would undermine (4').

Another way for Fisher to try to construct the required case would be to use a scenario in which almost no such interaction occurs – one in which the brains are not causally influenced by external events, so that they do not diverge. Perhaps, for example, Edna and Paula are meditating, and neither is engaging in any significant bodily activity or receiving significant environmental stimulation.

However, it then becomes far less clear that Edna and Paula are wholly mentally distinct, as (5') claims. Even assuming (what might be doubted) that their behaviors and situations still count as radically distinct, as (4') requires, it becomes unclear that those behaviors and situations are significant in individuating their mental states – given their much-reduced causal interaction with their brains. Remember that Fisher is concerned not just with mental content, but with the entire mental landscape – including, for instance, emotion-types. Thus if Edna is having happy thoughts about puppies, then Fisher wishes to claim not only that Paula might not be thinking about *puppies*, but that her thoughts might not even be *happy*. For the sake of argument, grant that her thought content could be something other than puppies. But could she be experiencing an emotion other than happiness from t_1 - t_2 ? Especially given the stipulated paucity of causal influence from their external features, I find this not plausible at all. Surely gross emotional experiences like happiness are determined by contemporaneous physical events in the subject's brain. Certainly I see no argument from Fisher to support a claim to the contrary. Yet the onus is on him to provide that argument.

Let me sum up the last few paragraphs. To run the revised argument, concerning a case in which the neural co-incidence extends beyond a moment, Fisher would need a case that meets three criteria. Firstly, as per premise (1'), Edna's and Paula's brains must be indistinguishable from t_1 - t_2 . Secondly, as per (4'), the external features must be notably distinct. Thirdly, to ground (5'), those features must be relevant to the individuation of the subjects' mental lives from t_1 - t_2 . The problem is that the first and second criteria tend to conflict, given the constant causal interaction between brains, bodies and environments.⁶ This problem *might* be overcome by describing a scenario that minimizes that interaction – but only at the cost of failing the third criterion. The less interaction there is between a brain and its surroundings through a given period, the less its surroundings individuate the subject's occurrent mental life in that period.

In fact, I think that (5') is problematic *no matter how distinct* Edna's and Paula's external features are. In the original argument, (5) appears cogent because a *momentary* co-incidence of brain states offers no resistance to the idea that two individuals who are as distinct as Edna and Paula must be distinct in all of their mental states. The co-incidence seems insufficient to determine any mental feature, so we are inclined to discount it and judge the subjects mentally distinct by virtue of their distinctness in all other respects. But an extended co-incidence is not so easily discounted, even if it occurs within the context of the sort of external distinctness claimed by (4'). We judge (5') to be false precisely because of cases in which, as per (1'), the neural co-incidence is at least length λ . Edna and Paula will plausibly have indistinguishable experiential states, which includes phenomenal experiences, emotion-types, and perhaps also occurrent propositional-attitude-types.

To drive the point home, consider an extreme case. Imagine two individuals, Alf and Abe, whose brains remain indistinguishable for an *hour*. It is, I take it, almost impossible to imagine that Alf and Abe could be significantly distinct in their bodies, activities and environments for

⁶ In the original argument the conflict did not occur because the neural indistinguishability only lasted for a moment, leaving no time for the causal effect of external features to spoil it.

that hour – so long as there is any causal interaction at all between their brains and those other features! But suppose, for the sake of argument, that those external features are pretty distinct. Even then, while there may be *some* mental differences between Alf and Abe, surely they are not *entirely* mentally distinct during that hour. The brain has far too central a role in our mental lives for that to be true. In particular, it is not at all plausible that they are wholly distinct in their experiences or in any experientially-tinged occurrent states. We would fully expect them to have the same phenomenal experiences and emotion-types.

That is an extreme case. But similar remarks apply, I think, if the neural co-incidence lasts for just a minute, or even for just a second. Only if the co-incidence is merely *momentary* – or at any rate, short enough to exclude the occurrence of determinate experiences – does the claim of radical mental distinctness become plausible. Of course, that plausibility disappears again if the snapshot conception of phenomenal experience is false, for I showed in §4 that we then need not suppose that there *are* any experiences precisely at the moment of neural co-incidence. So without the snapshot conception, Fisher’s argument becomes unsound. But I also showed that if the conception is embraced then the internalist may argue that it is only to be expected that the neural co-incidence would produce a phenomenal co-incidence, however odd that may seem. Either way, Fisher fails to show that an individual’s phenomenal experiences are not ‘just in the head’.

References

- Bartlett, G. (2008). Whither internalism? How internalists should respond to the extended mind hypothesis. *Metaphilosophy*, 39, 163-184.
- Burge, T. (1979). Individualism and the mental. In P. A. French, T. E. Uehling, and H. K. Wettstein (eds.), *Midwest Studies in Philosophy IV: Studies in Metaphysics* (pp. 73-121). Minneapolis: University of Minnesota Press.
- Burge, T. (1986). Individualism and psychology. *The Philosophical Review*, 95, 3-45.

- Clark, A. (2009). Spreading the joy? Why the machinery of consciousness is (probably) still in the head. *Mind*, *118*, 963-993.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*, 7-19.
- Fisher, J. C. (2007). Why nothing mental is just in the head. *Noûs*, *41*, 318-334.
- Hurley, S. L. (1998). *Consciousness in action*. Cambridge, MA: Harvard University Press
- McLaughlin, B., & Bartlett, G. (2004). Have Noë and Thompson cast doubt on the neural correlates of consciousness program? *Journal of Consciousness Studies*, *11*, 56-67.
- Noë, A. (2004). *Action in perception*. Cambridge, MA: MIT Press.
- Noë, A. (2006). Experience without the head. In T. S. Gendler & J. Hawthorne (eds.), *Perceptual experience* (pp. 411-433). Oxford: Oxford University Press.
- Noë, A., & Thompson, E. (2004). Are there neural correlates of consciousness? *Journal of Consciousness Studies*, *11*, 3-28.
- Putnam, H. (1975). The meaning of 'meaning'. In K. Gunderson (ed.), *Minnesota Studies in the Philosophy of Science VII: Language, mind, and knowledge* (pp. 131-193). Minneapolis: University of Minnesota Press.