

# Experiencing Left and Right in a Non-Orientable World

Jonathan Simon

Draft 11/20/18

Imagine that the person you see in the mirror were a real person, with her own experiences, living in an environment that is the mirror reverse of yours. Would her experience be the phenomenal duplicate of yours, or would her experience be different, the mirror reverse of yours? In other words, is the mirror-reverse of the neural correlate of your experience  $E$  also a neural correlate of  $E$ , or is it the neural correlate of a mirror reversal of  $E$ ?

Intuitively, your mirror twin's experience would be the mirror reverse of yours. If you are looking at your right hand, she is looking at her left hand. Plausibly, then, you are having an experience as of a right hand, while she is having an experience as of a left hand, and these are different experiences, the mirror reverse of each other.

Lee (2006) argues that this intuition is unsustainable if relationalism about space-time is true. For if relationalism is true, your twin's world is physically the duplicate of yours, and given physicalism, this means her experience must be the duplicate of

yours as well. Lee goes on to claim that relationalism has some very surprising consequences, for example: if our brains and bodies were perfectly symmetric we'd be unable to have asymmetric experiences. So we can deduce (transcendentally) from the asymmetric character of our experience that our brains or bodies are asymmetric.

But Lee's reasoning can be resisted. It is true that, if you and your mirror twin live in different possible worlds, then given relationalism, we can set things up so that your world and your twin's are duplicates, so that there can be no real physical sense in which you are incongruent, rather than congruent, with her. But matters are different if your mirror twin is your worldmate: say, your next-door neighbor, or your future self. The relationalist can maintain that mirror-reversed neural correlates (in mirror-reversed experiencers) yield mirror-reversed experiences *within* a given world. I'll call this the INTRAWORLD OPPOSITE EXPERIENCES principle. This allows the relationalist to resist Lee's striking transcendental deduction. So I argue in §1. In §2 I elaborate on what an INTRAWORLD OPPOSITE EXPERIENCES-consistent version of relationalism must look like.

This is not the end of the story, however. Whether or not relationalism is true, INTRAWORLD OPPOSITE EXPERIENCES cannot be a necessary truth. There will be some worlds for which it is not the case that mirror-reversed neural correlates (in mirror-reversed experiencers) yield mirror-reversed experiences. These will include worlds whose spatial topology is non-orientable (e.g., as a Möbius strip or Klein bottle are non-orientable). I argue for this in §3.

This is interesting, in part, because it means that one can draw something in the vicinity of Lee's transcendental deduction at those worlds. But there are also

ramifications for us back home. Relationalists who want to avoid Lee-style transcendental conclusions about the actual world are going to have to choose between dualism about the mental (the denial that the mental is grounded in the physical), a very radical externalism about the phenomenology of directional experience (the view that it constitutively depends on the overall topology of the universe), the view that the realization of directional experience is sensitive to asymmetries at the quantum level, and finally, the view that there is not a phenomenal difference between a total experiential state and its mirror reversal (a view Chalmers (forthcoming) dubs *e-relationism*). But rejecting relationalism does not get us out of the woods, either. I defend these claims in §§4-5. In §6 I conclude with the suggestion that e-relationism and dualism emerge as the most cohesive overall solutions to the puzzles considered.

## 1 Mirrors Between Worlds, Mirrors Within Worlds

Lee (2006) considers the following case. At the possible world  $w_1$ , Righty is looking at a sign that says ‘MIT’. The possible world  $w_2$  is the mirror reflection of  $w_1$ . Accordingly, Righty’s counterpart at  $w_2$  is Lefty, and Lefty is looking at a sign that says ‘TIM’.

Now, you’d expect that Right and Lefty would have distinct experiences. Let’s not oversell the point: you might also expect Righty and Lefty to say different things when asked what their signs say, but they wouldn’t, since they are molecule-for-molecule mirror reflections of one another, and mirror reflection doesn’t affect what sounds you make. So if Righty is going to utter the word ‘MIT’, Lefty will too.

Presumably, if Righty reads English from left to right, Lefty reads it from right to left.

However, as Lee (2006) points out, whether this is so depends on the metaphysics of spacetime. In particular, it depends on whether, beyond spatiotemporal relations like distance, betweenness, congruence, and so on, there are also absolute spatiotemporal features — for example, points of spacetime that can be re-identified from one world to another — and on whether, if there are any such further features, they can help determine what we experience.

Many dispute the existence of such further features,<sup>1</sup> and even those who countenance them for reasons pertaining to the philosophy of physics might shy away from taking them to figure in the determination of our experience. I'll call the view that denies that any such features figure in the determination of our experience, remaining neutral about whether they exist at all, *weak relationalism*, and I'll call the view that they do exist and figure in our experience, *strong absolutism*.<sup>2</sup>

If weak relationalism is correct, then  $w_1$  and  $w_2$  are physical duplicates of one another in every respect that matters for the determination of experience. This means that if the mental supervenes on the physical,  $w_1$  and  $w_2$  must be mental duplicates as well.

This is a sound point, but it leaves open the possibility of an *intra*-world mirror inversion case. What if Lefty and Righty are worldmates? For example, they might

---

<sup>1</sup>See e.g. Baker (2010), Belot (2012) and Dasgupta (2015). See e.g. Maudlin (2012) *pro contra*.

<sup>2</sup>As Baker (2011) notes, standard characterizations of the relationalism/substantivalism debate are not drawn in quite this way. Note for example that sophisticated substantivalism, which countenances regions of spacetime and their geometric relations in the ontology, but denies the re-identifiability of points or regions across worlds, would count as a form of weak relationalism in my sense.

be next-door neighbors, with a plane of mirror symmetry running along their shared property line. Alternatively, Lefty might be Righty's future self: the person he will one day be, after the mad scientists are done with him. And we cannot conclude, from the equivalence of  $w_1$  and  $w_2$ , that worldmate Lefty and Righty must be experientially equivalent, any more than we can conclude, from the equivalence of a world containing only a left glove, and a world containing only a right glove, that a world containing both a right glove and a left glove is equivalent to a world containing two left gloves. Where these intra-world scenarios are concerned, it is perfectly consistent with weak relationalism that Righty's experience is the mirror-reverse of Lefty's.<sup>3</sup>

Or is it? Here I want to consider a few observations made by Lee (2006) which suggest otherwise. Lee observes that, given weak relationalism, the result of mirror reversing Righty, but keeping the rest of the world the same, must be equivalent to the result of mirror reversing the rest of the world, but keeping Righty the same. But if a certain kind of change to Righty is equivalent to only changing Righty's environment and not changing Righty at all, it can't be relevant to Righty's phenomenology until he interacts with that environment. So then the difference between Righty and Lefty cannot amount to a difference in phenomenology (as long as they remain perfect mirror reflections of one another).

Lee notes an intriguing corollary, which is that a perfectly symmetric experiencer

---

<sup>3</sup>I note that at the actual world, parity violations (asymmetries in the statistics of particle decay) would render a true, perfect mirror reflection of an actual individual, down to the quantum level, at best an unstable statistical aberration, at worst nomologically impossible. My starting position is that it would be a move of desperation to hold that these quantum asymmetries are relevant to psychophysics. Accordingly, at the outset I'll neglect these asymmetries, and when I speak of Righty and Lefty being perfect mirror reflections of one another, I mean that they are perfect mirror reflections modulo parity violations and the like. I will revisit the issue below and consider the possibility that parity violations are relevant to the determination of directional experience.

could not notice if you mirror-reversed everything in her environment, because, by our previous reasoning, mirror-reversing her environment and not her is the same as mirror-reversing her but not her environment (as far as her phenomenology is concerned, given weak relationalism). However, since she is symmetric, mirror-reversing her changes nothing at all. Lee concludes that if weak relationalism is true, symmetric experiencers can only have symmetric experiences.

But this reasoning still depends on an *interworld* version of the thought experiment. If the scenario in question is intraworld — e.g., Righty’s life unfolding over time — then there *is* a difference between a future in which Righty is mirror reversed but the environment stays the same, and a future in which Righty stays the same but his environment is mirror reversed. Assuming the reversal is achieved through some physically possible motion, it will presumably involve rotation or other forms of acceleration, and any self-respecting relationalist owes an account of the effects of acceleration (i.e., inertial effects).<sup>4</sup> And this does not hinge on whether our experiencer is asymmetric. Say we have some perfectly mirror-symmetric experiencer; call her Ava. Assuming we can re-identify Ava’s parts over time, we can trace their paths across spacetime, and say whether they have all been reversed around, or whether they have remained inert while the world reversed.

However, there is a datum here which appears to have ramifications even on an intraworld construal of the thought experiment. Plausibly Righty’s functional, physical responses, if he is reversed and his environment is not, will be equivalent to those he would give if his environment had reversed and he had not. So we

---

<sup>4</sup>For discussion of relationalist approaches to inertial effects see, e.g., Dasgupta (2015).

can predict that either way, when he opens his eyes, he'll say "*The sign flipped! Before, it said 'MIT' but now it says 'TIM'!*" And this suggests that nothing changes phenomenologically for Righty until he opens his eyes and looks at the sign again.

But let's fill in some of the details. First, suppose that, before Righty is reversed, he closes his eyes and begins to imaginatively imagine the 'MIT' sign he was just looking at. He keeps his eyes closed, and continues to imagine this sign during the reversal. Now we ask: has his imagistic phenomenology been reversed, after he is reversed but before he opens his eyes? The datum about what he'll *say* might suggest otherwise. But in fact, either option is compatible with that datum.

The question is how Righty's brain realizes directional experience, and in particular, whether the mirror-reversal of the neural correlate of Righty's 'MIT' experience is itself the neural correlate of an 'MIT' experience, or is instead the neural correlate of a 'TIM' experience.

One part of the picture is how Righty's sensory apparatus is wired in with these neural correlates of his conscious states. Call the neural correlate of Righty's 'MIT' experience  $N_{MIT}$ , and call the mirror-reversal of  $N_{MIT}$ ,  $\overline{N_{MIT}}$ . Note that if some pattern of sensory stimulation  $S_{MIT}$  triggers  $N_{MIT}$  in Righty, the mirror reverse of that pattern,  $\overline{S_{MIT}}$ , will trigger  $\overline{N_{MIT}}$  in mirror-reversed Righty (i.e.,  $\overline{\text{Righty}}$ ).

Lee offers a colorful model of how Righty's sensory stimulation might encode directional information. Imagine that Righty is an English soccer enthusiast, with one half of his face (the left half) painted red, and the other half (the right half) painted white. Images that imprint on the red half of his face, he experiences as to the left, and images that imprint on the white half, he experiences as to the

right. Because his red half is his left half and his white half is his right half, when Righty looks square on at an ‘MIT’ sign, the ‘M’ imprints on his red half, and the ‘T’ imprints on his white half, with the ‘I’ squarely in the middle, so he experiences ‘MIT’. Here,  $S_{MIT}$  is a state of ‘M’ imprinted on red, and ‘T’ imprinted on white.

Crucially,  $\overline{S_{MIT}}$ , the mirror of  $S_{MIT}$ , is *also* a state of ‘M’ imprinted on red, and ‘T’ imprinted on white. But once Righty has been mirror-reversed into  $\overline{\text{Righty}}$ , it is the right half of his face that is red and the left half that is white. This means that after his reversal,  $\overline{\text{Righty}}$  would have to look at a sign that said ‘TIM’ in order to enter into the  $\overline{S_{MIT}}$  state again. An ‘MIT’ sign will now lead to a new imprintation pattern, ‘M’ on white, ‘T’ on red. This is a pattern that, before the reversal, would trigger the neural correlate of a ‘TIM’ sign experience,  $N_{TIM}$ , in Righty, and after the reversal, will trigger the mirror reversal of that neural correlate,  $\overline{N_{TIM}}$ , in  $\overline{\text{Righty}}$ .

With those details in place we can see how, even if  $\overline{\text{Righty}}$ ’s imaginative imagery is reversed when he is reversed (and before his eyes are open), he will say, “*The sign flipped! Before, it said ‘MIT’ but now it says ‘TIM’!*”. This is because, if  $\overline{\text{Righty}}$ ’s imaginative imagery is indeed reversed when he is reversed, so that  $\overline{N_{MIT}}$  realizes ‘TIM’ imagery, then  $\overline{N_{TIM}}$  realizes ‘MIT’ imagery. This means that when  $\overline{\text{Righty}}$  opens his eyes his perceptual phenomenology is of an ‘MIT’ sign but his imagistic memory phenomenology is of a ‘TIM’ sign.<sup>5</sup> And since he’s phenomenally reversed, he now reads from right to left, so his way of telling you that he now sees what you would call an ‘MIT’ sign is: “*I see a ‘TIM’ sign*” (if you’d asked him what he is imagining after the reversal but before he opened his eyes, he’d have said “*An ‘MIT’*”

---

<sup>5</sup>For ease of presentation I am assuming that his imaginative imagistic phenomenology is realized by the same mechanism as his perceptual phenomenology.

*sign*”, even though he’d have been imagining a ‘TIM’ sign).

I conclude that either option is consistent with the datum about what  $\overline{\text{Righty}}$  will say. It remains an open question whether the mirror-reverse of the neural correlate of an experience  $E$  is itself a neural correlate of  $E$ , or instead is the neural correlate of an experience which is the mirror-reversal of  $E$ . In other words, it remains an open question whether we *use* direction in order to represent it in experience, roughly as the indicator lights on your car use direction to represent it: your left blinker indicates that you’ll turn left; your right blinker indicates that you’ll turn right. Flip the blinker arrows around, and you flip what each of them represents.

We can operationalize this idea as follows: a neural correlate of an experience *uses direction to realize direction in experience* just in case the mirror reversal of that neural state (situated in a mirror-reversal of its host experiencer) is not a correlate of the same experience as the state itself.<sup>6</sup> I’ll call the theory that the neural correlates of our experiences use direction to realize it the *Use Theory*.

The key claim I’ve defended so far in this section is that weak relationalism does not automatically preclude the use theory, i.e. the theory that the neural correlates of our experiences use direction to realize direction. Of course, if weak relationalism is true, then a neural correlate and its mirror reversal are intrinsic duplicates in all respects relevant to the constitution of experience. So to countenance the use theory, one must be open to the possibility that experience is not entirely intrinsic.

---

<sup>6</sup>A terminological glitch: for ease of exposition I am neglecting micro-asymmetries for now, and considering mirror reversal modulo parity violations and the like. But intuitively, it should not count as using direction to realize direction if one’s neural states use alignment with respect to micro-asymmetries, rather than direction itself, to realize experience of direction. The reader may take me to be operating, for now, under the supposition that micro-asymmetries are not relevant to psychophysics. I revisit this issue below.

But many allow that experience is not entirely intrinsic.<sup>7</sup> It is also worth noting that relations to absolute regions of spacetime don't obviously count as intrinsic, so if there's a problem here, weak relationalism isn't obviously the culprit.<sup>8</sup>

Now, I say the use theory is open to the (weak) relationalist. But what exactly can we say about the grounds of directional experience, or the laws of experience, according to the use theory, given relationalism? I'll explore this in the next section.

## 2 The Psychophysics of Directional Experience

Let's develop a toy example to work with, to flesh out how the neural correlates of our experience might use direction to realize it. Suppose our subject is going to experience an arrow, pointing either to the right or the left. Conceivably, the neural correlate of such an experience might include an actual physical copy of the arrow (a constellation made out of spiking neurons in the shape of an arrow). The neural correlate of such an experience then uses direction to realize it in experience, if the actual direction in which this neural arrow points is what determines the direction that the arrow is experienced as pointing.

Let's think about all of the possible ways that directional experience might be realized (given either relationalism or absolutism). Here is a four-way partition, assuming that there is a phenomenal difference between an experience as of a left arrow and an experience as of a right arrow:

---

<sup>7</sup>Among others, Hurley (1998), Lycan (2001), Millikan (1989), Noë (2005), and Tye (2000). See Hawthorne (2004) *pro contra*.

<sup>8</sup>See Skow (2007).

**Symmetric Realizers** The experience of a left arrow is realized by a symmetric structure: e.g., the subject's experience as of an arrow pointing to the left is realized by a constellation of spiking neurons shaped like a circle (or, more realistically, by some computation based on a rate code or a temporal code whose spatial organization is inessential).

**Internally Asymmetric Realizers** The experience of a left arrow depends on the orientation of some neural processing with respect to an *internal* asymmetry as a reference point: e.g., the subject's experience as of an arrow pointing to the left is realized by a constellation of spiking neurons in the shape of an arrow pointing in the direction of the red side of her face.

**Externally Asymmetric Realizers** The experience of a left arrow depends on the orientation of some neural processing with respect to an *external* reference point, say, a glove in a vault at the National Institute of Standards and Technology: e.g., the subject's experience as of an arrow pointing to the left is realized by a constellation of spiking neurons in the shape of an arrow pointing in the direction that the thumb of the NIST glove would point if she held it palm inward.

***De Re* Direction Involving Realizers** : The experience of a left arrow depends on the orientation of some neural processing with respect to the directions themselves: e.g., the subject's experience as of an arrow pointing to the left is realized by a constellation of spiking neurons in the shape of an arrow pointing in *this* direction.

The use theory requires either externally asymmetric realizers or *de re* direction involving realizers. If you take the mirror reverse of a symmetric or internally asymmetric realizer in the mirror reverse of a subject, you'll realize the same experience: a mirror-reversed circle is still a circle, the mirror reversal of an arrow pointing toward the red side is an arrow pointing to the red side. But if you take the mirror reverse of an externally asymmetric or *de re* direction involving realizer, you'll realize a different experience (assuming that you don't mirror reverse the external reference object).

If relationalism is true, none of these options is going to be a picnic. Externally asymmetric realizer accounts are on their face the least plausible: why should it be this glove in Washington rather than that glove in France which sets the standard? There is a challenge of arbitrariness here, but also a challenge of radical externalism, or action at a distance. Our best hope might be to appeal to quantum parity violations — asymmetries in how certain particles tend to decay. Here we might appeal to the asymmetric law itself, rather than any specific instance.<sup>9</sup> However, it would be a surprise to learn that our experience depends in this way on the details of quantum mechanics.<sup>10</sup>

But symmetric realizer and internally asymmetric realizer accounts are also troubled. How can some symmetric bit of processing underwrite one orientation for experience rather than another? It would be one thing if we could claim, for example, that this bit of processing gives rise to Leftish experience because it is causally

---

<sup>9</sup>Does this help us avoid radical externalism? This depends on one's theory of laws, according to Hawthorne (2004). But see Weatherson (2005).

<sup>10</sup>It is also hard to classify this approach, since such a nomological asymmetry would be pervasively internal as well as external.

related to left-handedness, but if relationalism is true nothing is causally related to left-handedness, or anyway, no such relationships figure in the explanation of experience (I'll use 'leftish' and 'rightish' to designate mirror-reversed phenomenal characters, and 'left-handedness' and 'right-handedness' to designate the strong absolutist's directional properties<sup>11</sup>). Likewise for internal asymmetry accounts: why should the red side record what one experiences as leftish, rather than what one experiences as rightish?

Perhaps worse, symmetric and internal asymmetric realizer accounts also support something in the vicinity of Lee (2006)'s transcendental deduction. Obviously, if the realizers of leftish and rightish experience require internal asymmetries, then perfectly symmetric experiencers can't have them. The case of symmetric realizers is more subtle. There is no metaphysical bar on a symmetric experiencer's brain realizing such an experience, but there is a causal question of how she might manage to enter into one reliably in response to a stimulus.<sup>12</sup> If you have to get your neurons into a circle to experience a left arrow, and you have to get them into a diamond to experience a right arrow, then to be reliable there'll have to be some asymmetry at your transducers: e.g., those associated with the left side of your visual field might be wired to lead to the circular state, while those associated with the right side might be wired to lead to the diamond state. Left-side receptors and right-side receptors must then have been given different marching orders. But such an asymmetry in wiring runs against the assumption of true (temporary) symmetry.

---

<sup>11</sup>Obviously there is no general way of partitioning experiences into leftish and rightish ones. But the terminology is still helpful.

<sup>12</sup>See the related point in Baker (2011).

However, the right kind of sensory substitution system could allow such an experiencer to learn to see direction reliably. Such a system would need to translate asymmetric visual information into some symmetric format — this could include audition (i.e., linguistic descriptions), tactition, or indeed vision. Given a sufficiently neuroplastic brain, the subject could potentially attune her perceptual experiential states to these inputs. Arguably, if this happens, the sensory substitution system becomes part of her body, making her asymmetric after all. But arguably not.

Is there an alternative? That is, can the relationalist appeal to *de re* direction involving realizers? I say yes. After all, the relationalist must countenance orientational structure at a world. There are three questions to distinguish. Take the 2-D case for ease of exposition. First there is the *local* question of what it is for there to be two different rotational directions in which something might orbit around a given point on a surface. Second there is the *global* question of whether the surface is orientable, which amounts to a question of whether directions defined locally can be coordinated with one another, globally, in a smooth way.<sup>13</sup> Third there is a transworld question, of whether directions defined locally at points or regions of one possible world can be coordinated with those defined at another possible world. The relationalist should deny the possibility of such transworld coordination, but retain the possibility of intra-world coordination (on pain of sacrificing the distinction between worlds that are orientable and those that are not).

It is this account to which the relationalist wishing to avoid Lee-style transcendental conclusions should probably aspire. There is still a challenge of arbitrariness:

---

<sup>13</sup>See a standard topology text, e.g. Milnor (1965), for exposition.

since the directions the relationalist defines locally cannot be identified with absolute left-handedness or absolute right-handedness, there is still a question of why alignment in *this* direction should yield rightish experience rather than leftish. Worse, there are presumably as many realizers of a given directional experience as there are possible worlds, since given relationalism *this* direction at  $w_1$  cannot be automatically identified with *this* direction at  $w_2$ , even if they qualitatively have a lot in common.

This may be objectionable, but it does not seem to be decisive. So I conclude provisionally that a relationalist use theory is tenable. It remains possible for a (temporarily) perfectly mirror-symmetric Ava to reliably experience asymmetric things (like arrows), if the neural correlates of her experience use direction to realize it in her experience: her receptors can be symmetrically wired to record what is in front of them, meaning that a left arrow triggers a left-pointing neural assembly, and a right arrow triggers a right-pointing neural assembly, and given the use theory, the direction the neural assembly points can suffice to determine whether Ava experiences a left arrow or a right arrow.<sup>14</sup>

Second, it remains open that someone who was recently a mirror twin of yours still sees everything with the same orientation that you see it (though her memories differ from yours). That is to say, if you and someone who was recently your perfect mirror twin both begin looking at an ‘MIT’ sign, it is possible that you both still have the same ‘MIT’ experience. She will *say* that she sees a ‘TIM’ sign, but if the

---

<sup>14</sup>Presumably, once the designated assembly is pointing right rather than left, Ava is no longer perfectly symmetric. Lee’s question is whether an experiencer who temporarily enters into a perfectly symmetric state can be capable of asymmetric experience shortly thereafter.

use theory is true, she sees an ‘MIT’ sign just like you, though now she reads from right to left.

Despite this, proponents of the use theory shouldn’t get too comfortable. In the next section, I will develop an independent challenge to it. The challenge will show that the use theory can’t be correct on every possible world — even if we embrace strong absolutism. It will show that for the use theory to be correct on our world, we must either embrace dualism or a very disjunctive, and possibly radically externalist, conception of the grounds of conscious experience — even given strong absolutism.

### **3 Life on a 3-D Klein Bottle**

In the previous sections, I have argued that simply assuming weak relationalism isn’t enough to defeat the use theory. You can be a weak relationalist and still countenance the possibility that mirror-reversed twins have mirror reversed experiences. Here, however, I’m going to use a slightly more sophisticated thought experiment to show that there is a problem with the use theory after all, and indeed, that it affects absolutists as well as relationalists.

Iris and Siri are mirror-opposites, living near one another in mirror opposite environments at the same world. Say, both live in rooms on space ships, and these rooms are the mirror-reverses of each other. Iris has a tattoo on her left hand, and she is looking at it. Siri has a tattoo on her right hand, and she is looking at it. Assume, per the use theory, that their experiences are mirror-opposites: Iris is having an experience of a tattooed left hand, while Siri is having an experience of a tattooed

right hand.

It turns out that the universe in which Iris and Siri live is non-orientable, like a Möbius strip, but in three-dimensions. This means that Iris can travel around a curve in space, and end up back where she was, but mirror-reversed.<sup>15</sup>

Say Iris takes just such a journey, while continuously looking at her tattooed hand. Meanwhile, Siri remains stationary. Later, having completed the journey, Iris brings her ship back around to Siri's neighborhood. Now, Iris (and her ship) is a perfect duplicate of Siri (and her ship). Any passerby will tell you she sees two identical twins in two identical ships, both with tattooed right hands.

So now, when Iris looks at her tattooed hand, what does she see? Iris and Siri are now perfect twins, so they had better be seeing the same thing. But there is a big problem here. I stipulated that on her journey, Iris never takes her eyes off of her tattooed hand: she stares at it continuously the whole time. Ask Iris, and she'll tell you her experience remained constant the whole time. The problem is that (given that their experiences were distinct in the first place) we can't both say that Iris' experience remains the same at every point of her journey, and that her experience is the same as Siri's at the end of her journey. But we have good reason to think that both of those things are true. Hence, their experiences can't have been distinct in the first place.

Let's go through the options more carefully. If Iris' experience changes, is the change gradual, or sudden? It cannot be gradual. There are no degrees of leftness

---

<sup>15</sup>If you're uncomfortable with the 3-D version, you can get the idea by supposing that Iris and Siri are flatlanders living on a true Möbius strip or Klein bottle. For the actual version, various topologies suffice: for example, the product space of the projective plane cross the circle ( $\mathbb{R}P^2 \times S^1$ ).

or rightness for an (internally rigid) hand to change between.

But on the other hand, it is also not plausible that Iris' experience changes suddenly. At which point does it do so? There is no privileged boundary in this universe, such that one's handedness flips over the moment one crosses it. The processes underlying her journey are completely gradual. So there is no physical basis for any sudden reversal of experience. And Iris won't undergo any relevant intrinsic changes, so she certainly won't report experiencing any such reversal.

Why can't we just say that Iris' experience remains different from Siri's, even after her journey is complete? This is to say that they have different experiences even though you can put Iris right next to Siri, and see them to be perfect intrinsic duplicates, *and also congruent*, differing from one another only by a linear spatial translation, both with tattoos on the same hand, both with brains processing light reflected off those tattoos in precisely the same way (say their perceptual inputs are controlled within their spaceships, keeping them in perfectly duplicated states).

So maintaining that their experiences are different violates the most basic precept of physicalism, namely that duplicate physical states engender duplicate phenomenal states. Since we are assuming the use theory (for reductio), we are allowing that incongruent counterparts can have different phenomenal states even if they are intrinsic duplicates. But here we've got two *congruent* intrinsic duplicates of one another.

Admittedly, we haven't filled in enough details to guarantee perfect overall duplication, even at a time: their environments might differ *outside* of their pods, and their histories will presumably differ. However, it is hard to imagine a halfway plausible

theory of psychophysical correlations that will exploit these differences successfully.

Nor is it any help to opt out of physicalism at this point in the dialectic. It is hard to imagine a halfway plausible story of psychophysical correlations that accounts for how these two perfect duplicate worldmates differ experientially, whether those correlations are reductive grounding / identity claims (as the physicalist says), or nomological laws (as the dualist should say).

As it happens, it also does not help to jettison weak relationalism at this point in the dialectic. Even if there are objective properties of left-handedness and right-handedness, re-identifiable across worlds where they are instantiated, that doesn't help at non-orientable worlds, because these properties are uninstantiated are non-orientable worlds.<sup>16</sup>

I conclude that Iris' and Siri's experiences have been the same — phenomenal duplicates — for as long as they have been incongruent counterparts of one another. INTRAWORLD OPPOSITE EXPERIENCES is false at their world. In the next section (§4), I'll discuss what might be true at their world instead.

## 4 The Upshot for Iris and Siri

A question we have thus far avoided is whether there is a phenomenal difference between an experience as of a tattooed left hand and an experience as of a tattooed

---

<sup>16</sup>Or gluttily over-instantiated. It seems a minimal constraint on Absolute left-handedness and right-handedness, that if some direction instantiates left-handedness at every point in some region diffeomorphic with an open ball in  $\mathbb{E}^3$ , then the same direction instantiates left-handedness at every point in overlapping regions diffeomorphic with open balls in  $\mathbb{E}^3$ . But then on a non-orientable manifold both directions at every region will instantiate both left-handedness and right-handedness, if any directions at any regions instantiate either.

right hand, and in general, whether there is a phenomenal difference between a total experience with asymmetric content, and its mirror reverse. Chalmers (forthcoming) dubs the view that there is such a difference *e-categoricism* ('e' for experience) and the view that there is no such difference *e-relationism*. According to e-relationism, the differences between, e.g., an experience of a left arrow, and an experience of a right arrow, rest entirely in the relations between experiences. Thus far, we've been taking e-categoricism for granted. It might seem that e-relationism is patently incompatible with the phenomenology: obviously an arrow pointing left looks different than an arrow pointing right! But e-relationism is a thesis about one's *total* experience.

Note that the real difference in phenomenology at issue here is that between experiencing an arrow as pointing to *your* left and *your* right. This seems to include some awareness of the different parts of your own body. Hold that fixed, and vary the experience of which direction the arrow is pointing, relative to that fixed awareness of the locations of different parts of your body, and of course you'll get a different experience. E-relationism is consistent with that. The question is what happens when we reverse all of your experience of the location of these different parts of your body (visual, proprioceptive, tactile, etc) along with your experience of the direction of the arrow. Here, at least, the answer is not immediately obvious.<sup>17</sup>

Now, if e-relationism is true, that explains our result about Iris and Siri's world (the failure of INTRAWORLD OPPOSITE EXPERIENCES), and it implies that the same must

---

<sup>17</sup>See Chalmers (forthcoming) for a more extended discussion. For present purposes I am neglecting the distinction between e-relationism and what Lee (2006) dubs the Frege-Schlick view, according to which there is an intra-personal difference between leftish and rightish experience but no interpersonal one.

hold at our world. This speaks strongly in favor of e-relationism.

If we reject e-relationism, our options for understanding Iris and Siri's directional experience are limited. Either the realizers of Iris and Siri's leftish and rightish experiences are symmetric, or they are internally asymmetric. Either way, a qualified version of Lee's transcendental deduction can be drawn. Intriguingly, this is so even if we assume strong absolutism about our own world, since as we note just above, left-handedness and right-handedness are uninstantiated at non-orientable worlds.

This means that the cosmological discovery that our cosmos were non-orientable would put you in a position to transcendently deduce, from the introspection that your experience is asymmetric, that you yourself are physically asymmetric (or hooked up to a suitably asymmetric sensory substitution device).

Still, we have good reason to believe that our own universe is orientable. It is consistent with all I have said so far that the neural correlates of our directional experience use direction to realize it in our experiences, though the neural correlates of Iris and Siri's experiences do not.

But if we endorse reductionism about the mental, we aspire to a uniform analysis of phenomenal properties in physical terms. If pain is c-fiber firing at some other possible world, it is c-fiber firing at our world as well. Pain might be multiply realized: c-fiber firing in beings like us, d-fiber firing in other creatures — but here one hopes for some uniform demarcation at some other level, like the level of function or representation, which would account for the phenomenal identity.

The problem is this. If we find it implausible that Lee-style transcendental deductions should be available at our own world, we must either embrace e-relationism,

or maintain that there is a major difference between the way directional experience is realized on Iris and Siri's world and the way it is realized on our world, a difference in light of which our realizers of directional experience can use direction, though Iris and Siri's cannot. But this threatens to render our account of directional phenomenal properties profoundly disjunctive. Worse yet, the basic difference between Iris and Siri's world and our own doesn't boil down to anything local. Instead, it boils down to a difference in global topology: the fact that there is no consistent scheme of orientations that can be imposed across all regions in the universe. But this suggests that the realizers of directional experience depend on the global topology of the universe.

Making matters worse, there may be ways of bringing the puzzle of Iris and Siri closer to home. Earlier, we spoke casually of Righty and Lefty being flipped around, without considering how it might happen (building, as we were, on a discussion of *interworld* scenarios where the issue doesn't arise). But what generates our puzzle isn't the exotic nature of non-orientable manifolds *per se*, it is the way you can implement mirror-reversal gradually, on a non-orientable manifold, without disrupting the flow of experience. And this can be done on an orientable manifold if it has extra spatial dimensions. Just as you can pick up a 'b' on a piece of paper and put it back down as a 'd', so you might rotate through an extra spatial dimension and come back mirror-reversed. And there are prominent physical theories on which there are, indeed, more than three spatial dimensions, so we may have a harder time distancing ourselves from an Iris and Siri scenario so developed.

In the next section (§5), I'll review the options if we want to differentiate our world

from Iris' and Siri's, and maintain that at our world, mirror opposite experiencers would have mirror opposite experiences, even though this isn't true at worlds like Iris and Siri's. Most of these options have significant downsides, and the only one that allows us to maintain a truly uniform analysis of the phenomenal properties involved is property dualism. Finally in §6 I will summarize the master argument.

## 5 Bringing the Problem Home

There are various ways that we might differentiate our world from Iris and Siri's, in order to maintain the INTRAWORLD OPPOSITE EXPERIENCES intuition at our world, despite its falsity at theirs. Here I assess them.

Suppose we wish to hang on to weak relationalism as well as physicalism about experience. Then we must identify some difference between the circumstances of Iris and Siri on a non-orientable world (or higher-dimensional world) and an analogous pair of characters on our own world, a difference suitable to *ground* the difference in phenomenology that keeps non-orientable/higher-dimensional Iris and Siri from having opposite experiences but allows actual Iris and Siri to do so.

Apparently, the only way to do so is by appealing to the global topology itself. That is, we must count the global topology of the universe among the grounds of any single instance of a directional experience (because it is only thanks to our universe's being orientable that, say, the core realizer of a leftish experience is externally asymmetric or *de re* direction involving).

What is so bad about that? Above, we observed that for the use theory to be

tenable in a relationalist world, we'll have to deny that perfect *intrinsic* duplicates are phenomenal duplicates. But to many, that won't come as news. Numerous arguments have been given for why consciousness is not perfectly intrinsic, depending on, for example, whether a thing is detached or undetached,<sup>18</sup> and on various features of its environment and history.<sup>19</sup>

But there's extrinsic, and then there's extrinsic. It is one thing for an entity's consciousness to depend on its local environment. It is another thing for it to depend on the entire global spatial structure of the cosmos. But if we were to claim that experience of left and right somehow plays out differently for experiencers who live on non-orientable manifolds, we'd be claiming precisely that. As I've stressed, orientation can be defined locally, meaning that, barring the considerations raised here, we might construe *de re* direction involving realizers as local, though it is a global matter whether all of the local definitions cohere.

It is hard to conceive of a motivated alternative, or indeed of any alternative that preserves the use theory. At this point, one might be tempted to appeal to quantum parity violations: actual asymmetries in the statistics of particle decay. Because a non-orientable world presumably cannot (stably) have asymmetric decay laws, we might take these to be constitutively involved in our realizations of directional experience, but not in Iris and Siri's.

On this picture, for example, one might experience an arrow as pointing to the left if the neural correlate of that experience points in the direction in which beta-particles tend to decay, and as pointing to the right if the correlate of that experience

---

<sup>18</sup>Sider (1998).

<sup>19</sup>Hurley (1998), Lycan (2001), Millikan (1989), Noë (2005), Tye (2000).

points in the opposite direction.

As I note above, it is hard to classify this picture: is it a version of use theory, or not? As I note above, in a world (like ours) with asymmetries built into the laws, it is not possible for two beings to (stably) coexist which are strict mirror reversals of one another. So use theory at worlds like ours is better understood at the macro-level: whether beings who are reversed at the macro level have reversed experiences or not. So construed, the picture in which parity asymmetries are what determine the leftishness or rightishness of experience is consonant with the use theory.

Still, we generally expect to find the correlates of experience at the neural, or at worst the chemical level. It would be shocking indeed to find that our neural states depended in this striking way on the quantum state. And of course, even though this approach might allow us to resist the charge of radical externalism concerning our own realizers of directional experience, we still wind up with a disjunctive theory of the relevant phenomenal properties. We'd still need an entirely different sort of dependency to explain how directional experience arises at Iris and Siri's world, and to explain how it arises at worlds that are orientable, but happen to have no asymmetries at the quantum level.

An alternative is to reject weak relationalism in favor of strong absolutism. We saw that absolute left-handedness and right-handedness cannot be instantiated at non-orientable worlds. But one might offer this up as a rationale for why Iris and Siri are experientially different from us. That is, one might take it that the directional experience we enjoy constitutively involves absolute direction, and accordingly that Iris and Siri, living on a non-oriented manifold, must not have the same kind of

experience that we have. This solution feels less ad-hoc than one appealing to parity violations at the quantum level. It allows us to invoke *de re* direction involving realizers of directional experience at our own world with far more systematicity than the weak relationalist can hope for, since the realizers can make reference to left-handedness and right-handedness themselves, rather than merely *this direction* and *that direction*.

But even this approach leaves something to be desired, for it still leaves us in the dark about the nature of Iris' and Siri's experience at non-orientable worlds.<sup>20</sup> So the account remains disjunctive: we get one form of realizer in worlds where left-handedness and right-handedness are instantiated; an entirely different form in worlds where they are not instantiated.

Another alternative is to deny that the phenomenal properties instantiated by Iris and Siri are distinct from those instantiated at our world by people looking at tattooed hands. Iris and Siri's phenomenology might be alien, in which case there need be nothing disjunctive about the use theoretic analysis of our directional phenomenology. But it is obscure what this alternative phenomenology would be (as Lee 2006 and Chalmers (forthcoming) have noted). Second of all, the spectre of radical externalism still looms. In virtue of what, precisely, does Iris' neural state realize an alien phenomenal experience rather than the sort that you or I experience? Unless we bring parity violations or absolute directions back into the picture, nothing local differentiates Iris from you or me, so it appears to continue to depend on global topology.

---

<sup>20</sup>Perhaps it covers the case of higher-dimensional orientable worlds, since the two directions at the 3-D hypersurface on which Iris and Siri live could instantiate left-handedness and right-handedness.

We have one final account to consider of how it might be that INTRAWORLD OPPOSITE EXPERIENCES is true at our world though it is false at Iris and Siri's. This is property dualism.

The property dualist can maintain that phenomenal properties are intrinsic, and that each of them has a non-disjunctive essence, while allowing that the psychophysical laws that determine who instantiates them are dependent on such things as the global topology of the universe. And where it is implausible to imagine that *what it is* to experience an arrow pointing to the left is some relation to global topology, it is perhaps not as difficult to imagine that the laws that specify the links between the phenomenal and the physical should be attuned to such things.

Lee (2006) and Chalmers (forthcoming) note another putative case for dualism in the neighborhood. Consider the trans-world case again: Righty at  $w_1$  and Lefty at  $w_2$  are both having the same experience. But assuming e-categorialism, this means there are two experience types, 'MIT' and 'TIM', such that necessarily one is uninstantiated in all worlds that physically duplicate  $w_1$  and  $w_2$ . But given the symmetry between these experience types, and given relationalism, it is hard to see how there could be any difference between their modal status.

This argument can be pressed into the mold of standard explanatory gap arguments. For example, it might be couched as a spin on the complaint that physicalism commits us to brute necessities, necessities that violate Chalmers' principle that 1-conceivability entails 1-possibility.<sup>21</sup>

Arguably, however, there is more to the present argument. Here, there is indis-

---

<sup>21</sup>Chalmers (2009).

putably a perfect symmetry between the two candidates. There is something particularly compelling to the charge that metaphysical possibility shouldn't discriminate between perfectly symmetric scenarios. Despite this, neither Lee nor, surprisingly, Chalmers, take dualism to be the proper moral (Chalmers takes the moral to be e-relationism).

According to Chalmers, there is a problem here for the dualist who wishes to preserve the nomological supervenience of the mental on the physical. For if the 'MIT' experience and the 'TIM' experience are both nomologically possible, given a pair of worlds that are physical duplicates, nomological supervenience is violated.

But this is not a compelling objection to dualism. One can insist that metaphysical possibility not discriminate between perfectly symmetric scenarios while allowing that nomological possibility do so. Perhaps the 'MIT'-experience scenario just turns out to be nomologically possible while the 'TIM'-experience scenario turns out to be a world with different psychophysical laws.

Nor is the dualist out of options who insists on allowing that both possibilities are nomological possibilities. The dualist can qualify the commitment to nomological supervenience, say, by allowing that the psychophysical laws are indeterministic or probabilistic, and permit two different 'phenomenal orientations' consistent with any given physical configuration. Here, the dualist resembles the strong absolutist, who countenances worlds that share the same material geometry but differ by boosts in absolute velocity or shifts in absolute location. But the dualist has the advantage, because she only countenances metaphysical differences that make a phenomenal difference.

## 6 The Master Argument

So here is where things stand. You can embrace e-relationism. It solves all of the problems we've considered. It explains why INTRAWORLD OPPOSITE EXPERIENCES must be false at every possible world, while avoiding all arbitrariness that seems to arise if we say that a given experiencer's experience is leftish rather than rightish. It allows us to treat directional experience at Iris and Siri's world in precisely the same way as we treat our own, while also allowing us to resist Lee-style transcendental deductions, and remain neutral on the mechanisms by which directional experience (such as it is) is realized.

But if you just can't bring yourself to embrace e-relationism, something has to give. You must then allow that one can transcendently deduce one's physical asymmetry from introspection, given that one's world is non-orientable. And then you must choose between accepting that the same is true even if one's world is orientable, accepting dualism, and accepting a range of analyses of experience that render it disjunctive and worse. The options are:<sup>22</sup>

---

<sup>22</sup>I take the option wherein Iris and Siri have alien phenomenology to be subsumed under global topology externalism, and I consider the Frege-Schlick view to be a species of e-relationism.

1. **e-relationism**
2. **Transcendental Deductions at Home**
3. **Disjunctive Analyses**
  - (a) **Global Topology Externalism**
  - (b) **Quantum Mind**
  - (c) **Strong Absolutism**
4. **Property Dualism**

Typically, it is the scientifically-minded philosophers who eschew dualism who hold that *a priori* philosophy of mind should not ‘interfere’ with the honest work of the scientists. Here, ironically, it is precisely the hard-headed scientific mindset which leads the philosopher to dictate to the scientist from the armchair: if consciousness is a (non-disjunctive) natural kind, and dualism is false, then the neural correlates of our directional experiences can’t use direction to realize it in our experience, and the realizers of our leftish and rightish experience must themselves be either symmetric or internally asymmetric, meaning that symmetric experiencers cannot have leftish or rightish experience.

The philosopher who wishes to maintain a hands-off approach, while also maintaining the non-disjunctiveness of specific experience-types, must therefore choose between e-relationalism and property dualism. And again, this is not just a predicament for weak relationalists, since strong absolutism cannot help Iris and Siri. <sup>23</sup>

---

<sup>23</sup>My thanks to David Chalmers, Brian Cutter, Dan Korman, Dana Goswick, Kristie Miller and Jennifer Wang for helpful comments, not all of which have yet been incorporated into this draft.

## References

- Baker, D. (2010). Symmetry and the Metaphysics of Physics. *Philosophy Compass* 5 (2010): 1157-66.
- Baker, D. (2011). The experience of left and right meets the physics of left and right. *Noûs* 46:483-498.
- Belot, B. (2012). *Geometric Possibility*. Oxford: OUP.
- Chalmers, D.J. (2009). The Two-Dimensional Argument Against Materialism. In Brian P. McLaughlin & Sven Walter (eds.), *Oxford Handbook to the Philosophy of Mind*. Oxford University Press.
- Chalmers, D.J. (forthcoming). 'Three Puzzles about Spatial Experience'. (Pautz and Stoljar, eds.) *Themes from Ned Block*. Oxford: OUP.
- Hawthorne, J. (2004). Why Humeans Are Out of Their Minds. *Noûs* 38 (2):351-58.
- Hurley, S. (1998). *Consciousness in Action*. Oxford University Press..
- Lee, G. (2006). The experience of left and right. In (T. Gendler and J. Hawthorne, eds.) *Perceptual Experience*. Oxford University Press.
- Lycan, B. (2001). The Case for Phenomenal Externalism. *Philosophical Perspectives* 15 (15):17-35.
- Maudlin, T. (2012). *Philosophy of Physics: Space and Time*. Princeton, NJ: Princeton University Press.
- Millikan, R. (1989). Biosemantics. *Journal of Philosophy* 86 (July):281-97.
- Milnor, J. (1965). *Topology from a Differentiable Viewpoint*. University Press of Virginia. Charlottesville.
- Noë, A. (2005). *Action in Perception*. MIT Press.
- Skow, B. (2007). Are Shapes Intrinsic? *Philosophical Studies* 133: 111-130.
- Tye, M. (2000). *Consciousness, Color and Content*. MIT Press.
- Weatherson, B. (2005). Humeans Aren't Out of Their Minds. *Noûs* 41 (3):529-535.