# Experiencing Left and Right in a Non-Orientable World

Jonathan Simon

## 1 Introduction

Imagine that the person you see through the looking glass is a real person, with her own experiences, living in an environment that is the mirror-reverse of yours. You look at your right-hand glove as you put it on your right hand; she looks at her left-hand glove as she puts it on her left hand. You feel your heart beating on your left side; she feels her heart beating on her right side.

Is there any difference between your experience and hers? She is looking at her left-hand glove, the mirror-reverse of you as you look at your right-hand glove. Does she see what you see, a right-hand glove, or the reverse of what you see, a left-hand glove?

Since her environment is the mirror-reverse of yours, and she seems to get by in it just fine, you'd probably assume, if you had to guess, that she sees it correctly. She is *looking* at her left-hand glove, after all. So it stands to reason that she is having an experience as of a left-hand glove, not a right-hand glove. And it stands to reason

— it seems obvious — that an experience as of a left-hand glove differs in what it is like from an experience as of a right-hand glove.

In this paper I will argue that this last presupposition, obvious though it may seem, is false. I'll argue that there is no intrinsic difference between what it is like to experience a left-hand glove and what it is like to experience a right-hand glove: if there are any relevant experiential differences, these are either merely comparative (differences in what it is like to *compare* gloves) or cognitive (differences in what it is like to bring an experience of a glove under a concept like LEFT or RIGHT). So your mirror-reverse twin, who won't differ from you in what she is comparing or thinking, will have the same experience as you. I'll call the position I am arguing against *orientation-experience categoricalism*, and the position I am arguing for *orientation-experience relationism*.[1]

Orientation-experience relationism is a surprising thesis, one whose falsity you might take for granted. I will argue for it by cases. Things go wrong in one sort of way if we imagine that you and your mirror-twin's experiences differ: basically, this leads to absurdity if we imagine that you and she both live at a world with a non-orientable spatial topology (in the sense in which a Möbius strip or Klein bottle are non-orientable). Things go wrong in a different sort of way if we assume that you and your twin would experience the same, but that someone else might have the experience that we were imagining above that your twin might have (i.e., we hang on to orientation-experience categoricalism). Under these assumptions we have to accept the existence of partial zombies, who lack the relevant phenomenology but

[1]I mean by these terms the same positions that Chalmers (2019) labels 'e-categoricalism' and 'e-relationism', restricted to the case of experience of orientation.

behave as if they have it, at very close, empirically possible worlds.

Why care? Well, the general question of how to reconcile our experience of left/right orientation with the nature of orientation itself has been with us since Kant. But the fact that we have to choose between orientation-experience relationism and the choice of poisions I'll outline is newsworthy.[2] Also, at least some of the symmetries that give rise to trouble in the case of orientational experience have analogues in other cases, like color experience, so even those who don't feel invested in the outcome of this debate may find that positions here have ramifications elsewhere (which I hope to explore in future work).

Here is the plan. In §2 I'll set the stage. In §3 I'll talk about what goes wrong if we assume that you and your twin have different experiences. In §4 I'll talk about what goes wrong if we assume that you and your twin have the same experience, but that someone else might still have the experience that in §3 we imagined your twin to have. In §5 I conclude with a discussion of orientation-experience relationism. I also include one appendix, which outlines what I think is the best fallback option if you appreciate the worries I raise in §3 and §4, but cannot bring yourself to reject orientation-experience categoricalism. This is to appeal to dualism to defend the view I attack in §3.

[2]Complementary arguments for orientation-experience relationism can be found in Lee (2006) and Chalmers (2019).

## 2 Setting the Stage

This is a paper about orientation: a feature of spaces, or objects in spaces. As is custom, I will make gloves my primary example. Your two gloves are not perfectly identical: if one of them fits on your left hand, the other does not. Your gloves are *incongruent counterparts*.

It is tempting to think of the features over which your gloves differ as part of, or anyway akin to, their shapes. Unfortunately, it isn't that simple. Shape might be intrinsic, even if relationalism about spacetime is true (to be a circle is for all surface points to be the same distance from a center point),[3] but orientation cannot be. Orientation is more like size: to say whether one circle is smaller or bigger than another, you have to invoke relations between the two (or take your distance relations to specify absolute magnitudes).

However, orientation is more problematic even than size. You can consider size to be an intrinsic, absolute feature if you want to. With orientation, you don't even have the option. Earman (1989, p.140-141) explains why left-handedness and right-handedness cannot be intrinsic properties of gloves (or the regions of space they occupy). If there were such a thing as intrinsic left-handedness, nothing on a non-orientable world could have it — even if spacetime is a substance rather than a network of relations (I'll elaborate in §3 below). But a left-handed glove on our world could be a perfect (intrinsic) duplicate of a glove on a non-orientable world. Since intrinsic duplicates share all intrinsic properties, it follows that there's no relevant intrinsic property for gloves around here, either. In other words, the thesis we might

[3]Though see Skow (2007).

call *spatial-orientation categoricalism* is false.[4]

The funny thing is that it sure *seems* like there is an intrinsic difference between right-handedness and left-handedness. It certainly seems as though you can tell the difference (and not just that there is one) simply by looking. It is not as though all you can do is say, of two gloves, whether they are congruent or incongruent. Shown a single glove, there seems to be an intrinsic *way it looks* that justifies your classifying it as left-handed, much as you can look at an apple and see that it is red, without having to compare it to a green apple in the process. That there is such an intrinsic way it looks is the claim that I am calling *orientation-experience categoricalism*. If you deny this claim, and accept what I am calling *orientation-experience relationism*, then you have to deny that a world with just a single experience of a left-hand glove is any different from a world with just a single experience of a right-hand glove, though you can make sense of some associated differences at more amply furnished worlds.

This paper is an extended argument for orientation-experience relationism. Again, I don't claim that this thesis is intuitively plausible; it is just that the alternatives seem to be worse. Ultimately, the lesson is that orientation-experience categorical-

[4]Cutter (2019) offers a simpler argument. Step one: no property that you can lose or gain just by being slid around — i.e., by rigid translation — is plausibly intrinsic. Step two: given a space with the right topology: for example, a space with an extra dimension, you can change a left-hand glove into a right-hand glove just by sliding it around, much as you can change a 'p' into a 'q' by lifting it off the page. But here's an objection to step one. We can imagine a world where intrinsic properties (colors, say) come and go spontaneously. But then why not also intrinsic properties that change when you slide things around? I suspect that Cutter is ultimately moved by the sense that any change from intrinsic-left to intrinsic-right would be arbitrary. But there are ways the friend of intrinsics might minimize the arbitrariness. For example, maybe you can only be intrinsically a 'p' while on the page, so the letter token loses the property as soon as it is lifted, and then acquires intrinsic 'q'-ness as soon as it is put back down. As we will see in §3, Earman's appeal to non-orientable spaces closes this loophole.

ism goes together with spatial-orientation categoricalism, and the latter's being false makes the former hard to defend (even, as we will see, if we embrace a substantival theory of spacetime).

To consolidate: I've now discussed two distinct notions of intrinsicality. First, the notion of intrinsicality of handedness itself; second, the notion of intrinsicality of experiences of handedness. This can get confusing. Bear in mind that claims of intrinsicality need an index: we have to say what a given property is intrinsic *to*. In the case of spatial orientation (spatial-orientation categoricalism), the question is whether the glove's property of being left-handed is intrinsic to the glove itself. In the case of experiences of handedness (orientation-experience categoricalism), the question is whether there is an orientational phenomenal character of the experience of the glove (what we might call its *leftishness* or its *rightishness*) that is intrinsic to whatever it is that grounds the experience of the glove — even if these grounds of the experience are external to the body of the experiencer. Here, the question is whether you have to add some other experience of comparison or cognitive activity onto the experience of the glove in order to get anything we might call an experience of orientation.

To complicate matters, there is a *third* notion of intrinsicality that we'll need in order to frame the debate. This is the notion of the intrinsicality of experience to the experiencer, the idea that the facts about an experience all supervene on, and are fully metaphysically determined by, things that are inside the experiencer's brain, or anyway their brain and body, or anyway their brain and body and things in the immediate environment of that brain and body at the time in question. I'll call this

the *local supervenience* thesis.

Though the local supervenience thesis and orientation-experience categoricalism both concern the intrinsicality of experience, they are independent of each other. You can maintain local supervenience and deny that there is an intrinsic difference between leftish and rightish experience. You can also maintain that there are intrinsic differences between leftish and rightish experiences, while maintaining some form of externalism-to-the-experiencer about the grounds of that experience. This compares to a position like externalist representationalism about color,[5] according to which there is an intrinsic or categorical difference between an experience of red and an experience of green — these are inherently different "flavors" of phenomenology — but the grounds of such experience involve facts about the experiencer's evolutionary history.

Our original question about what your mirror-reverse twin would experience turns out to be pegged to whether or not local supervenience is true. This is because, as we have seen, spatial-orientation categoricalism (the view that left-handedness and right-handedness can be intrinsic properties of gloves) is false. From the falsity of spatial-orientation categoricalism, it follows that you and your mirror-reverse twin are intrinsic (and therefore local) duplicates. Thus if local supervenience is true, your experience must be the same as hers.

This means that to make sense of the idea that your experience differs from your mirror-reverse twin's, we have to reject local supervenience. How might this go? One very wild idea is that somehow the spatial incongruency relation between you and

[5]E.g., Tye (2000).

your twin directly determines the difference between your phenomenology and hers. A slightly less wild idea is that there is some reference object, say, a statue of a glove next to the original *mètre étalon* in Paris, such that whether you experience a leftish or a rightish glove depends somehow on a congruency relation between something on your end and that statue. An even more promising idea, though one that will take more work to spell out, is that the orientational character of your experience is a function of how you embed into spacetime. As I will explain, this idea can be motivated (and made coherent), even if spatial-orientation categoricalism is false (and also even if substantivalism about spacetime is false).

So we can reframe the argument to follow in terms of local supervenience. In the next section, §3, I'll be considering the case in which local supervenience is false (but orientation-experience categoricalism is true). We can then make sense of the idea that your mirror-reverse twin experiences something different from you, a left-handed glove where you see a right-handed one. Then in §4 I'll consider the case in which local supervenience is true (and orientation-experience categoricalism is too). This entails that your mirror-reverse twin experiences the same thing as you (because she is your perfect local duplicate) though other experiencers who are not your perfect local duplicates might still experience leftish where you experience rightish.

# 3 Denying Local Supervenience

## 3.1 How Your Twin's Experience Might Differ From Yours

If we want your twin's experience to be different from yours, we'll have to reject local supervenience. But it is not enough to reject local supervenience; we also need a positive account of how non-local elements contribute to determining you and your twin's experience (and the difference between them).

We can distinguish between *singular* and *pervasive* approaches. Singular approaches appeal to some singular (or anyway restricted) class of external reference objects. Pervasive approaches appeal to something more pervasive, like a quantum field, or the collection of all gloves in the universe.

Say we've got a character named Righty, looking at his right-hand glove, and having an accordingly *rightish* experience (remember, that's our word for one of the two distinctive flavours of categorical, orientational phenomenology that the orientation-experience categoricalist posits). Let's suppose that his brain's contribution to the realization of this experience has an orientational structure, say, an assembly of neurons arising in response to the stimulus, in the shape of a glove.

A singularist account might appeal to some specific object, external to Righty, like a statue of a glove somewhere in Paris. On such an account we might say that what determines that Righty has a rightish rather than a leftish experience is that the glove-shaped neural assembly in his brain is (physically, spatially) congruent with the designated glove-statue in Paris.

This account is, of course, wildly implausible. Say that the statue is a statue of

a right-hand glove. If a sculptor in Paris transforms it into a left-hand glove, will Righty's experience immediately change (because now his internal state is incongruent with the glove)? That sounds like action at a distance. Also, how on earth could this particular statue come to be linked to Righty this way, though he may be on another continent? And even granting that, why does congruency with this magic statue yield rightish rather than leftish experience?

A step in the direction of a pervasivist account solves some of these problems. Suppose for example that every glove can play the relevant role: the active one in a given episode of experience is the one that causes Righty's glove-shaped neural assembly during that episode (i.e., the glove he is looking at). Thus some gloves, the ones we call 'right-handed', are such that he has rightish experience iff his glove-shaped neural assembly is congruent with, and caused by, one of them, while other gloves, the ones we call 'left-handed', are such that he has leftish experience iff his glove-shaped neural assembly is congruent with, and caused by, one of them.

Note that if such a theory were correct we could use it to give a response-dependent theory of 'categorical' left- and right-handedness (though presumably the properties involved, constitutively involving relations to experiencers, are not really categorical). But bear in mind that since spatial-orientation categoricalism is false, we cannot think of the intrinsic orientation of the glove as explaining why congruency with it yields rightish rather than leftish experience. Things would rather be the other way around: the facts about which flavor of experience congruency with a given glove gives rise to would ground the facts about the glove's 'categorical' handedness.

Problems remain. Suppose you could (magically, say) mirror-reflect such a glove, without Righty seeing it. Then you'd be able to reverse his experience from rightish to leftish, without altering his internal state in any salient way (suppose, for example, he's just looked at the glove, but then you ask him to close his eyes and continue to sustain the image of the glove in iconic memory, and you do your magic while his eyes are closed: the view we're considering entails that his iconic memory experience would shift when you cast your spell). This would again be action at a distance (though admittedly a shorter distance, unless he's in Paris). Then there are the vagaries of causation to consider; these will lead to further trouble for the view.

Fully pervasive views solve these problems. A fully pervasive view takes orientational experience to be a function of how Righty's internal state (i.e. his glove-shaped neural-assembly) relates to local pervasive features, like a designated quantum field, or spacetime itself.

It is actually well-known that the weak interaction has an orientation in the statistics of particle decay. For example, a certain pattern of lambda baryon decay only occurs when a quark called the beauty quark has a counter-clockwise spin in its direction of motion (LHCb Collab, 2015). So we might say that Righty has a rightish glove experience just when he has a glove-shaped neural assembly that is congruent with the direction that beauty quarks spin when they generate this pattern.[6]

A closer look reveals that this appeal to quantum subtleties is unnecessary. Any network of objects will do for which local congruency relations are defined. Imagine,

[6]Say, align the fingers of the neural "glove" in the direction of the particle's linear velocity and cup the particle in the "palm": then congruency here can mean that the particle spins in the direction the "thumb" curls inward.

for example, that the world were pervaded with lots of very small gloves, both left- and right-handed ones. Local congruency — congruency relative to region or neighborhood of spacetime $n$ — between two gloves is then the question of whether they can be superposed onto one another via rigid translation alone without leaving $n$. We then have a partition of the gloves in $n$ into two equivalence classes, A and B. We can think of these as determining experience directly. For example, we might say that Righty has a rightish glove experience just when he has a glove-shaped neural assembly that is congruent with the gloves in (and so a member of) A. The picture here is that the bundle of local congruence relations between members of A taken together amount to the realizer of Righty's experience (with his glove-shaped neural assembly being the core of that realizer).

On this account, in effect, we can think of the principles of the psychophysics of orientation as *choices of phenomenal orientation* for neighborhoods of spacetime. Crucially, we don't really have to posit lots of little gloves (or oriented quantum fields) for the account to work. In the appendix I will explain how to do without them, and discuss the account's costs and benefits in greater detail. One clear benefit: the account allows us to say that Righty's experience will be reversed if he, along with everything in his neighborhood, is mirror-reversed from one moment to the next. For then given some neighborhood of spacetime that spans the period of this transformation, each of the gloves on the $t_1$ end will be incongruent with their future counterparts on the $t_2$ end.

The account requires that we constrain overlapping neighborhoods to agree with one another in the obvious way. If $n_1$ and $n_2$ overlap, their congruency equivalence

classes will overlap. If the equivalence classes of $n_1$ are A and B, and the equivalence classes of $n_2$ are C and D, then A will share members with C, or with D, but not both. Assume A shares members with C. Then we may require that if congruency with A determines rightish experience in $n_1$, then congruency with C determines rightish experience in $n_2$. If not, there is no consistent experience rule for things in the intersection of $n_1$ and $n_2$.

Incidentally, we can use this little-gloves construction to approximate what it is for a space to be globally orientable: it is for it to be possible to non-trivially extend the local equivalence classes defined on neighborhoods into a partition of two distinct global equivalence classes. In the case of a non-orientable manifold like a Möbius strip, this condition fails. An equivalent formulation: a space (suffused with little gloves) is globally orientable just in case there is a coherent choice of rule of orientational experience (i.e. one that respects the constraint I describe in the previous paragraph).

## 3.2 The Problem: Non-Orientable Worlds

This brings me to the central argument against views that reject local supervenience and assert that mirror-reverse twins have mirror-reversed experiences. Say that Iris and Siri are mirror-reverse twins, 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. We'll assume for reductio that their experiences are opposite one another's: Iris is having

a leftish experience of a tattooed left hand, while Siri is having a rightish 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 spatial dimensions. This means that Iris can maintain a statuesque rigid pose and travel around a curve in space, and yet end up back where she was, but mirror-reversed. If you find this hard to imagine, you can get the idea by supposing that Iris and Siri are two-dimensional flatlanders living on a proper Möbius strip or Klein bottle. Imagine that you have a translucent Möbius strip and a letter on it written in black ink that looks like a 'p' viewed from above, and thus like a 'q' viewed from below. Say the letter is written on a little cell of plastic, so that you can slide it around. Then you can slide this letter around the surface of the strip so that when it is next in the same spot, it will look like a 'q' from above and a 'p' from below. In effect, Möbius strips only have one side.[7]

So say that 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. Because of the properties of non-orientable space, Iris and her ship are now perfect duplicates of Siri and her ship: that is, they are intrinsic duplicates, as they were at the outset, but now in addition they are locally congruent. Any passerby will tell you she sees two identical twins in two identical ships right next to each other, both with tattooed right hands. By travelling around a twisted loop in her universe, Iris has mirror-reflected herself.

[7]For discussion see Gardner (1982). A three-dimensional example is the product space of the projective plane cross the circle ($\mathbb{RP}^2 \times \mathbb{S}^1$).

From this we can conclude that if at the outset Iris' experience differed from Siri's (our assumption for reductio), then there will be some moment at which two locally congruent intrinsic duplicates differ in what they experience. For either Iris' experience has changed so that on her return it is the same as Siri's (though when she left it was not), or it has not changed, so that on her return it remains different from Siri's.

If it changed, it did not happen gradually: there are no intermediary *shades* of orientational experience between leftish and rightish. So if her experience changed it happened abruptly, between one moment and the next. But that means there is some time-step, some $t_1$ and $t_2$ such that the $t_1$ stage of Iris has leftish experience and the $t_2$ stage of her has rightish experience, even though these stages are very close to each other in spacetime (and the spatial regions she was passing through at the relevant times are very close to each other in space). Because of those facts of proximity, we can still make sense of assertions of local congruency: Iris at $t_1$ is locally congruent with Iris at $t_2$. Thus we have two locally congruent intrinsic duplicates who differ in what they experience.[8]

Suppose on the other hand that Iris's experience does not change. Then at $t_{final}$, once she is back from her journey, standing next to Siri, Iris and Siri will be locally congruent intrinsic duplicates who differ in what they experience. Thus, either way, we'll have two locally congruent intrinsic duplicates who differ in what

[8]Incidentally, this is the step of the argument that makes our appeal to non-orientable space necessary. On the version where we imagine Iris being lifted up into a fourth spatial dimension, flipped over and then put back down (the way you or I might lift a 'p' off of the page and then put it back down a 'q'), 3D congruency relations are not well-defined transversally to the surface Iris lives on (just as there's no facts about whether two cut-out letters p/q that you are holding in your hand are congruent or incongruent) — orientation and congruency are top-dimensional behaviors.

they experience.

This is bad. Of course, we are considering views that reject local supervenience. Two locally congruent intrinsic duplicates might be related to different things: they might have different evolutionary histories, for example, or different glove statues might have impressed upon them at some critical developmental window.

But we can stipulate that Iris and Siri don't differ in those ways, and of course we know already that Iris at $t_1$ and Iris at $t_2$ do not differ in those ways. It follows that the only salient distinction between Iris and Siri at the outset is their local incongruence. Remove that local incongruence, and there is nothing left to ground any prospective difference in their experience. In other words, allow that Iris at $t_1$ differs experientially from Iris at $t_2$ (because Iris' experience changed), or that Iris at $t_{final}$ differs experientially from Siri at $t_{final}$ (because Iris' experience did not change), and you're stuck with *groundless* differences in experience between locally congruent intrinsic duplicates: differences that only a dualist could make sense of.

We can put this another way: there are no relevant qualitative (i.e., non-haecceitistic) differences suitable to ground any difference in leftish and rightish phenomenal quality between Iris at $t_1$ and Iris at $t_2$ or between Iris at $t_{final}$ and Siri at $t_{final}$. There's a shaky limb to go out on here: argue that merely haecceitistic differences can somehow ground qualitative ones. But if that's possible, I'm not sure what 'grounding' means anymore.[9]

[9]C.f. Pautz (forthcoming). Note that this reasoning also tells against responses which introduce some buffer zone in which, between experience leftish and experiencing rightish, Iris experiences neither, and this is a partial zombie. Partial zombies of this sort will be relevant in §4 below. But such a zombie between $t_1$ and $t_2$ would equally violate our principle of sameness of experience for locally congruent intrinsic duplicates, so it is not an option for the materialist who rejects local supervenience (though of course if it were, it would be a problematic reply, for the reasons I will

Ultimately, in the appendix, I will argue that a version of dualism emerges as the best option for those who can't part with orientation-experience categoricalism. But the version I recommend will be distinct from the one that we'd be stuck with if we insisted on experiential differences between locally congruent intrinsic duplicates here. The psychophysical laws required for that would be ugly indeed. What happens exactly, in regions where intrinsic duplicates (with the same evolutionary history, etc.), congruent within that region, can differ experientially?

The most promising way to reject local supervenience, as I suggest above, construes the rules for rightish and leftish as choices of phenomenal orientation, in a way that parallels the definition of spatial orientation (where locally an orientation is a choice of congruency equivalence class, and then global orientation is a question of whether each regions two equivalence classes can be stitched together into two distinct, coherent global ones). But if locally congruent intrinsic duplicates can differ experientially, then however experience is nomologically determined in regions where this happens, it won't amount to anything like a choice of phenomenal orientation. Thus the arbitrariness that confronts the dualist here is especially pernicious, because it requires that we abandon the parallel between phenomenal and spatial orientation that best motivates the approach (of rejecting local supervenience). In contrast, the version of dualism I will tentatively suggest as our second best option in the appendix will not reject local supervenience at non-orientable worlds relevantly like ours.[10]

outline in §4 below.

[10]Compare remarks in Lee (2006) and Chalmers (2019) about the problems for dualist accounts of orientational experience given relationalism about spacetime. Chalmers for example observes that nomologically possible relationalist worlds with just a single experiencer looking at a glove

I conclude that there is no sensible way to maintain that mirror-reverse twins differ experientially, at least at non-orientable worlds. This leaves us with two options (that preserve orientation-experience categoricalism). Either we embrace an account that respects local supervenience necessarily, so that Iris and Siri's orientational experiences have the same kind of realizers that ours do, or we embrace a disjunctive account, so that Iris and Siri's experiences are realized one way and ours are realized another. In §4 just below I'll review the problems with the first option. I'll cover the problems with the second option in §5 and the appendix. My conclusion will be that our best bet is to reject orientation-experience categoricalism.

A final observation: now that we've talked about non-orientable spaces we are in a position to understand Earman's argument that there can be no such thing as intrinsic right-handedness (for gloves). If there were such a property, then it would have to be shared by any two locally congruent gloves, but also, no single glove (at a single time) can be both intrinsically right-handed and intrinsically left-handed. Now imagine a world like Iris and Siri's, but with gloves everywhere. Suppose some glove, Iris-glove, is intrinsically left-handed, and some other nearby, locally incongruent glove, Siri-glove, is intrisically right-handed. No one has to travel, but we can trace a line of gloves around the space connected by a chain of local congruency relations, beginning with Iris-glove. Every glove in this chain must be intrinsically left-handed if Iris-glove is. But follow the chain around and we'll end up back where we started, with a glove that is locally congruent with Siri-glove. That glove must be intrinsically

would have to all be rightish or all be leftish. But one might reply that the relevant nomological principles are probabilistic. This would come with complications, but it could still cohere with the thought that, at a given world, the laws still assign a coherent choice of phenomenal orientation. It is this latter prospect that is thwarted when we consider non-orientable worlds.

left-handed, because it is linked by the chain, but it must be intrinsically right-handed because it is locally incongruent with Iris-glove, and locally congruent with Siri-glove. Contradiction.

# 4 Preserving Local Supervenience

## 4.1 How Your Twin's Experience Might be the Same as Yours

If local supervenience is true (local, intrinsic duplicates have the same experiences), then your mirror-reverse twin has the same experiences as you. But if orientation-experience categoricalism is correct, there is still a difference between leftish experience and rightish experience (and presumably you can have either).

How might leftish and rightish experiences of gloves both be realizeable within you, in a manner that respects local supervenience, given that there can be no such thing as intrinsic right-handedness or intrinsic left-handedness?

A given realizer of a given experiential state will generally involve a combination of some stimulus-responsive occurrent processing (what your brain does because you saw what you did), and some standing structural or background features (what your brain was doing anyway, its neural weights and biases, the location of Broca's area, etc.). When we imagine a global mirror-reversal, the kind that distinguishes you from your mirror-twin, we imagine a reversal that reverses everything, including spatially relevant features of these standing structures, so for example if your Broca's area is in your left hemisphere it would be in the right-hemisphere of a mirror-reverse twin

of yours.

We can also hold all of the standing structures fixed, pick some specific pattern of occurrent, stimulus-response processing $p$, and ask what happens to you when instead of that pattern you realize its mirror-reflection — i.e., a second pattern $\overline{p}$ of occurrent stimulus-responsive processing that is the reflection of $p$ about your median sagittal plane (the middle of your brain).

Call two total brain states that are related in this way (standing structures the same, occurrent processing $p$ reflected about median sagittal plane with occurrent processing $\overline{p}$) *p-reflections* of one another. Note that in the event that the processing state $p$ is already symmetrical along your median sagittal plane, a brain state is its own p-reflection.

We can divide the option space in terms of what you experience when p-reflected, if $p$ is the occurrent stimulus-responsive processing that determines your leftish or rightish experience. Say that you have *internally symmetric realizers* if p-reflection leaves your experience the same. Say in contrast that you have *internally asymmetric realizers*, if p-reflection changes what you experience (though global mirror-reversal, Broca's area and all, does not). Rejecting local supervenience would open up a third option, according to which p-reflection and global mirror-reversal would both reverse your experience in the same way.

An example of a system of internally symmetric realizers: leftish and rightish experiences are realized by different temporal codes, implemented in some region along your median sagittal plane. An example of a system of internally asymmetric realizers: you have a leftish glove experience just in case the palm of your glove-

shaped neural assembly points behind you when its thumb points in the direction of your Broca's area when centered on your median sagital plane; a rightish glove experience if the opposite.

Lee (2006) and Baker (2011) observe that funny things happen when you take experiencers with orientational experience that respects local supervenience (thus, in my terminology, experiencers whose leftish and rightish experiences have internally symmetric or internally asymmetric realizers) and make them perfectly symmetric.[11] Things get weird in different ways depending on which case we consider (I distinguish the cases as I do in part to help clarify this). If the realizers of leftish and rightish experience are internally asymmetric, they hinge on some internal asymmetry: some asymmetric feature in your standing or background neural structure to measure your occurrent processing against. But then if you become fully symmetric, those features will vanish, so you cannot have leftish or rightish experiences anymore, full stop. On the other hand if your realizers are internally symmetric, then if you were wired in a perfectly symmetric way, seeing a left-hand glove on your left side would lead to a state that is the p-reflection of the state you would enter into upon seeing a right-hand glove on your right side. But then (by definition of internally symmetric realizers) it follows that you'd have the same experience either way. Thus you'd be unable to reliably distinguish left-hand gloves from right-hand gloves on the basis of your experience.

This implies something curious: if you know that local supervenience is true and that your experiences of leftish and rightish are systematically reliable, then you can

[11] See also Hurley (1998).

deduce that you contain some internal asymmetries. Perhaps this insight may be used to help explain why there are so many cognitive asymmetries.[12]

But this is not just a curiosity: it leads to trouble. We can characterize a cognitive architecture that would allow an experiencer to perceptually distinguish left from right, even if that experiencer were to become perfectly symmetrical, or close enough. The problem is that local supervenience together with orientation-experience categoricalism will entail that experiencers with this architecture are able to reliably distinguish left from right perceptually and behave accordingly, without the conscious experience that we supposedly rely on to do the same thing, even though the behavior in question has all of the hallmarks of conscious rather than unconscious perception. The situation given local supervenience together with orientation-experience categoricalism would be analogous to one in which a color-blind person accurately puts red apples and green apples in different piles, just by looking.

So local supervenience together with orientation-experience categoricalism will entail that the experiencers in question are partial zombies: their apparently conscious behavior is just like ours but the experience that explains that behavior in us is missing in them. This is inherently implausible, and it undercuts our reasons for positing that there is such experience in the first place.

[12]Consider for example our capacity to use the completely asymmetric words, 'left' and right' to designate symmetric things. There must be some asymmetry of wiring that decides which speech action to initiate, otherwise for any stimuli that results in you (correctly) saying 'it is on my left', the mirror reverse of that stimuli on the other side of your visual (or tactile, or whatever) field would result in you (incorrectly) saying the same thing.

## 4.2 The Problem: Symmetric Architectures and Partial Zombies

Consider the way that your car indicator lights work: you signal that you will turn left by flashing a light that is physically on the left side of the car. You signal that you will turn right by flashing a light that is physically on the right side of the car. This is an elegant arrangement, not least because it allows for many interchangeable parts: two interchangeable buttons lead by two interchangeable wires to two interchangeable lights, differentiated only because one set is on the left and the other is on the right.

Along the same lines, we might build a simple robot capable of differentiating stimuli coming from the left or from the right, and non-verbally expressing as much — just swap out the car light buttons for perceptual sensors and connect each light to the nearest arm, in such a way that the arm raises if the connected light is lit. Thus, the robot raises its left arm when it sees something on the left, and its right arm when it sees something on the right.

Now I ask, might there be a being with this sort of underlying architecture that can engage in this kind of performance consciously? That is, might there be a being, organic or artificial, that can consciously see a glove, and then use that conscious perception to guide an action like putting it on the correct hand, without this conscious process causally depending on any standing asymmetries?

I say yes. Though my argument doesn't hinge on it, I'll mention the actual human case to frame the issue. The human early visual system does appear to be wired in a symmetric way, with left visual hemifield information from both eyes going via the right optic nerve to the right hemisphere, and right visual hemifield information from

both eyes going via the left optic nerve to the left hemisphere. So our transducers themselves don't seem to break the symmetry: if there were standing asymmetries in the process they would have to be in the cortical regions of the hemispheres themselves. Of course, our brains are pervasively lateralized, and many functions like language comprehension[13], speech production[14] and facial recognition[15] do seem to depend on standing asymmetries, perhaps essentially.[16] However, none of these capacities are implicated in the behavior I just described (consciously putting on a glove), so even if they are somehow constitutively involved in the realization of conscious experience, they are not causally implicated in the relevant instances of conscious processing.[17]

If you have misgivings about that, note that our test case needn't be an actual human. For my purposes it suffices that there is some nearby possible world where a fluke of evolution (or a feat of engineering) leads to such a creature: a creature in whom no standing asymmetries play a causal role in consciously perceiving a glove and being guided by that percept in voluntarily putting the glove on the hand it fits. My hypothesis is not that such a creature can function normally in all respects without any lateralization; nor that consciousness is possible without any such lat-

[13]

[14]

[15]

[16]See above note **??**.

[17]Might some of them still be implicated in a roundabout way, say, as some kind of gate-keeper for entry into the global workspace, or becoming an object of attention or meta-cognition? If so, it is hard to see how their asymmetry itself might be an essential causal ingredient. Various conditions such as dyslexia, autism and schizopherenia are associated with higher degrees of brain symmetry, presumably without preventing consciousness, and one assumes that if split brain patients and other cases can be conscious with only one hemisphere (Volz and Gazzaniga 2017). Finally, various forms of aphasia to lateralized functions, like Broca and Wernicke aphasias, do not appear to preclude consciousness (Brookshire 20076).

eralization, but only that essentially lateralized mechanisms might not be causally implicated in processes like the described one of consciously putting on a glove. If they have a constitutive role to play in the determination of leftish and rightish experience, it would be that of a constitutive *background* or *anchoring* condition, rather than that of a core realizer. I take the state of what we know about lateralization in the human case to show that, if this is not true of humans, this is because of something fairly accidental, meaning that some fluke of evolution or feat of engineering might lead to a creature of which it is true. This is what I will presume below.

So now let us say we have an accordingly wired conscious creature at a nearby possible world, call her Aviva.[18] Aviva is perfectly symmetric as far as the ready states of her transducers and perceptual processors leading to simple consciously guided motor behavior are concerned — though not necessarily in other respects: her Broca's area might still be in her left hemisphere, and this might still be a constitutive background condition for her experience. We just require that it be causally independent or downstream of the relevant perceptual processes — and that she nevertheless can consciously put on a glove.

I am suggesting that we have good empirical reason to take this to be possible, because toy models of the relevant architecture are trivial to construct, and there is no evidence that the principles underwriting these toy models could not also undergird suitably sophisticated ones. I am now going to argue that if we must accept one of the accounts of leftish and rightish experience consistent with local supervenience, then Aviva's existence is highly problematic.

[18]Lee calls his perfectly symmetric experiencer 'simple Simon'. But (ahem) let's stick to palindromes. It's ok: I say so.

If local supervenience is true, the realizers of Aviva's leftish and rightish experience are internally symmetric or internally asymmetric. But by design, her occurrent processing when she sees a left hand glove (on her left) is the reflection of her occurrent processing when she sees a right hand glove (on her right). Thus if her realizers are internally symmetric, she has the same experience either way. This means that if her experiences of leftish and rightish help explain her behavior (she puts the glove on her left hand because she has a leftish rather than rightish experience of it) under ordinary circumstances, the realizers of this experience had better be internally asymmetric.

So we'll assume the realizers of her leftish and rightish experience are internally asymmetric. But again by design, whatever standing features of her brain that might make the difference here (so that reflecting them about her median sagittal plane along with her occurrent processing leaves her experience the same, but reflecting the occurrent processing without reflecting them makes for a distinct experience) must be causally independent (or downstream) of the process of her consciously putting on a glove.

This means that there is some asymmetric standing structure in her brain — say, for example, that it is the position of her Broca's area — which is such that both a) it is only in relation to it that her stimulus-responsive occurrent processing determines rightish or leftish experience and b) it is outside of the causal loop from perception to consciously putting on a glove.

Now suppose that Aviva develops aphasia in this structure (her Broca's area) and the neurons involved wither away. Even if the aphasia is so pervasive that she loses

all linguistic ability, still by hypothesis she can carry out the behavior of putting on a glove efficiently just by looking at it, in an apparently conscious, voluntary way. It is highly plausible that actual Broca or Wernicke aphasics are conscious (global aphasia, the combination of both, is sometimes induced by stroke, and is consistent with a range of ordinary conscious functionality[19]): all the more so Aviva, since we have explicitly postulated that the area of aphasia is not causally involved in basic conscious processing in her case.

But given what we are supposing, this aphasia wipes out an essential component of the would-be realizers of Aviva's leftish and rightish experience. In other words, aphasic Aviva would be unable to experience leftish or rightish (at least without having equal and opposite experiences of the other), and yet she would continue putting on her gloves, as if by conscious sight, just as before.

Now, many hold that perception and action both come in unconscious as well as conscious varieties, and there is much debate over what differentiates conscious from unconscious varieties.[20] But that is a debate over how to differentiate the functional or causal structure of one kind of process from another. In Aviva's case, the functional or causal structure of the process pre-aphasia is by hypothesis the same as the process post-aphasia.

I conclude that aphasic Aviva is a partial zombie: she exhibits a kind of performance that we exhibit on the basis of our experience of leftish and rightish (at least if orientation-experience categoricalism is correct), but she does not have that experience. We can assume that the underlying contents have entered into her global

[19]Brookshire (2007).
[20]Block (2007), Carruthers (2019).

workspace, are attended to, are the objects of meta-cognition, and so on. Still, in aphasic Aviva, they do not yield leftish and rightish experience.[21]

This is problematic, for two reasons. First, I take it that the conclusion that partial zombies walk amongst us, or at worlds very much like ours, is a black eye in its own right. Here, indeed, equipped with the right kind of equipment — say, a transcranial magnetic stimulation (TMS) machine — you could turn someone into a partial zombie just by pressing a button. Again, this wouldn't just be a matter of shutting off an experiential feed (you can do that by putting your hands over someone's eyes), but of shutting off an experential feed without interfering with the behavior that seems to happen on the basis of that experience!

Second, the situation is one in which leftish and rightish experiences have little to no explanatory role left to play. If aphasic Aviva can get by just fine without leftish and rightish phenomenology, why can't we? If there's no need to appeal to leftish and rightish phenomenology to explain how Aviva gets by, we could probably manage without it ourselves.[22] If so, we might as well accept orientation-experience relationism.

On this latter point, the source of the trouble is not just that we've allowed some things outside of the causal loop into the realizers of Aviva's experience. Many

[21]What if the internal asymmetry on which Aviva's leftish and rightish phenomenology depended were haecceitistic: a question, say, of whether the thumb of the neural-assembly glove points toward *this* neuron rather than *that* one, even though the two are functionally identical? This makes our problematic result even easier to come by: just imagine that the relevant neurons are lesioned out and replaced by other functionally identical ones. That said, as I suggest in the previous section, I'm not sure if it makes sense to propose that merely haecceitistic differences can ground qualitative ones: if there can be any basic constraints on what can ground what, this seems plausibly to figure among them, at least if we intend for grounding to capture a sense of reduction or metaphysical explanation of the grounded by the grounds (see again Pautz forthcoming).

[22]Compare Shoemaker (1982) and Lee (2018).

allow that, but still take experience to play an explanatory role. For example, the externalist representationalist can take the character of my experience (its being greenish rather than reddish) to help explain why I say that I see a green apple (rather than saying that I see a red one), even though this fact about the character of my experience may depend on my evolutionary history.

Of course externalists (who reject local supervenience) tend to be more at ease with this sort of thing than internalists (who reject local supervenience). This is certainly a sign that something is amiss: if you reject externalism (and so, insist on local supervenience) *because* you find this sort of thing unacceptable, you are in hot water now.

But it is worse than that. The factors that externalists typically invoke here, like evolutionary history, are not subject to intervention or manipulation, and so can be held fixed for explanatory purposes. You can't knock my evolutionary history offline with a TMS machine. It is the fact that interventions like this are possible in Aviva's case that precludes any explanatory role for her leftish and rightish experiences. So the friend of local supervenience trying to find an explanatory role for Aviva's leftish and rightish experience is substantially worse off than the typical externalist.

Note that if spatial-orientation categoricalism were true we wouldn't be in this mess. There'd be no problem if we could say that she experiences leftish when her neural glove is intrinsically left-handed and rightish when it is intrinsically right-handed. But since we can't say that, if we want to preserve both local supervenience and orientation-experience categoricalism, we have to appeal to causally irrelevant but nevertheless manipulable standing elements to account for leftish and rightish

experience, given an architecture like Aviva's.

I conclude that the combination of local supervenience and orientation-experience categoricalism is best avoided. If we're set on orientation-experience categoricalism, we should be prepared to reject local supervenience, at least at the nearest worlds with experiencers like Aviva, so that we can preserve her ability to experience leftish and rightish even should she suffer from Broca's aphasia.[23] Partial zombies might be inevitable if Aviva lives on a world like Iris and Siri's, but by rejecting local supervenience around here we can at least keep them at bay.

# 5 Orientation-Experience Relationism

## 5.1 Overview: The Case for Orientation-Experience Relationism

So this is where we are. If we defend orientation-experience categoricalism, our choice between accepting and rejecting local supervenience is a choice of poisons. If local supervenience is true at worlds like ours, experiencers like aphasic Aviva are partial zombies. If local supervenience is false, we run into big trouble at globally non-orientable worlds like Iris and Siri's.

[23]You might wonder how the reasoning given here about left and right generalizes to other quality spaces like color space, which also exhibit symmetries. There seems to be a general lesson about the problems that arise if some aphasia or surgery can disable experience without affecting function (of at least some paradigmatic perception-to-action pathways). The question for other experience types is what might compel one to accept that. Here, I've appealed to specific empirical intuitions about what kinds of architectures are feasible for the case of leftish and rightish experience, and in the previous section I made special use of the topology of non-orientable spaces. It remains to be seen what analogues of these moves there may be in other cases. For work along related lines see Lee (2016)'s discussion of the synchronic Bertha case, as well as Shoemaker (1982,2006).

We can try to thread the needle by saying that local supervenience is true at Iris and Siri's world but not at Aviva's. The main problem with this (setting aside that creatures like aphasic Aviva at Iris and Siri's world would be partial zombies) is that it leads to a very disjunctive conception of the realizers of leftish and rightish experience. We are accustomed to the idea, defended by most functionalists, that a given experiential quality is multiply realizable: a specific feeling of pain might be realized by c-fiber firing in me, but by something entirely different in an octopus. Crucially, the functionalist offers us some solace: all of the different realizations play a common functional role. Still, some argue that this is not solace enough, and that the functionalist cannot account for the intuition that phenomenal similarities are genuine objective similarities (see e.g. Block 2016 and Pautz 2017).

But if there is a problem for the functionalist, there is an even bigger problem here — that is, on the strategy where we maintain that local supervenience is false at worlds like ours (the closest worlds with beings like Aviva) but allow, as we must, that something else is true at non-orientable worlds. For to say this — at least if we eschew dualism — is to say that leftish experience is grounded in or identical to a disjunction of states that do not implement a common functional role (at least not in any ordinary sense of 'functional role'), and are united only insofar as they offer competing accounts of how leftish or rightish experiences might be realized.

One way of getting at the problem with this is to think about facts of objective similarity. There is patently a sense in which all instances of a given phenomenal quality (like a specific feeling of pain, looking scarlet red, or looking leftish) are more similar to each other than they are to anything else, or at least to some instance of a

different phenomenal quality. But how could Aviva's leftish be more similar to Iris' leftish than to, say, Aviva's rightish, or anything else, if the natures of their realizers are so radically different that one respects local supervenience and the other does not?

Worse yet, this disjunctivism won't quite solve all of our problems. Even if we pay that price, since we must accept local supervenience at non-orientable worlds like Iris and Siri's, an aphasic Aviva living on such a world would be doomed to partial zombiedom. The best we can therefore do, it seems, is to focus on keeping the partial zombies at bay. The line we would have to take is that we preserve an explanatory role for experience provided that partial zombies are not close empirical possibilities, things that some engineer might actually create in the actual future, or that might have evolved owing to an evolutionary branch at a world otherwise just like ours. As long as partial zombies are far enough away in possibility space, we might argue, perhaps they don't undercut the explanatory role of the relevant experience around here. In the appendix I will explore this proposal further. But suffice it to say, it is a predicament that we avoid if we reject orientation-experience categoricalism.

The moral is that no matter what, orientation-experience categoricalism comes at a high cost. We'll have to reject local supervenience for worlds like ours, but accept it at non-orientable worlds. This commits us to an especially pernicious sort of disjunctivism (unless dualism is true) as well as the possibility of partial zombies at non-orientable worlds. On the other hand, while rejecting orientation-experience categoricalism is certainly *prima facie* counterintuitive, it is hard to identify any tangible theoretical costs of doing so, so we should at least consider it.

## 5.2 Where To Go From Here: Varieties of Orientation-Experience Relationism

Orientation-experience relationism comes in two flavors: structuralist and cognitivist. According to structuralist varieties, though there are no categorical phenomenal features of leftish or rightish, still there are *comparative* phenomenal facts, such as the fact that this experience is opposite in orientation to that one. We must then decide whether these comparative facts can hold between experiencers, or only between distinct experiences of the same experiencer. The latter approach has been called the 'Frege-Geach' view.[24] A variant form of structuralism builds the relationality into the content of the experience, so that leftish and rightish experiences amount to assessments of which other things the percept is congruent with.[25]

According to cognitive varieties of orientation-experience relationism, the differences between leftish and rightish experiences as we know them are cognitive rather than perceptual. The cognitivist denies that there is a difference between how we typically experience left-handed and right-handed things in perceptual phenomonology, but affirms that there are associated cognitive differences — differences which may exploit the various lateralized mechanisms downstream of or independent from basic perceptual processing (like language processing). An extremist version of cognitivism denies that these cognitive differences make any difference to the phenomenology, while a more moderate version of cognitivism allows that these cognitive differences make phenomenal differences, because some cognition gives rise to cogni-

[24]See Shoemaker (2006), Lee (2006,2016).

[25]This seems to be the way to extend the perceptual structuralism defended in Morrison (2015,2020) to the present case.

tive phenomenology.[26] For example, the asymmetric language processing circuitry of our terms 'left' and 'right' may somehow figure in conscious thoughts that typically co-occur with perceptual experience of oriented objects, or it may figure in concepts that cognitively penetrate our perceptual experiences.

Crucially, the proposal here isn't that we identify the (categorical, proprietary) perceptual experience of leftish with some realizer that involves conceptual categorization, but that we point to some other more variable phenomenon, downstream of perceptual processing, that may co-occur with the target just enough to explain away our intuition of something categorical and proprietary in the perceptual phenomenology.

The cognitivist picture is that Aviva's perceptual phenomenology does not deliver her an experience of leftish rather than rightish, though if she has no aphasias, her capacity for thought or cognitive phenomenology do allow her to consciously add some meta-data to her experience, annotating it as 'left' rather than 'right'. On this view we deny that Aviva needs this meta-data (or a properly leftish experience) to use her experience as of a glove to guide her action of putting it on.

Not any way of filling in the details will do. Structuralism still allows us to raise the challenge about Iris and Siri. Instead of asking whether Iris' experience remains leftish or flips to rightish, we ask whether it remains congruent with what it was, or flips to become incongruent with what it was (and congruent with what Siri's is). This will give rise to the same problem that the original case does, if experiential congruence is to be pegged somehow to local spatial congruence of realizers.

[26] Pitt (2004).

But structuralism puts new options on the table. One option is to hold that congruency and incongruency relations only hold at close spatio-temporal range, denying their transitivity. Another is to take congruency facts to involve a hidden parameter, such as path dependence. Thus a given pair of states (say, Iris and Siri's) could be congruent relative to one path (say, a short one that doesn't leave the neighborhood) but incongruent relative to another (say, the path of Iris' journey around the universe). This would effectively be to endorse a gauge-theoretic conception of experience (c.f. Maudlin 2007).

Cognitivist varieties of orientation-experience relationism have a more transparent way of resolving all of our difficulties, since they deny that there is any difference in perceptual phenomenology between leftish and rightish experience. Thus cognitivists deny that as far as sensory or perceptual phenomenology is concerned, there is any difference, categorical or relational, between Righty and Lefty, Iris and Siri, or Aviva and the rest of us. And even so, by appealing to cognitive phenomenology they needn't go as far as to deny that left and right make any difference to our phenomenology. This view also gives a seemingly fitting account of what happens to Aviva (post-aphasia): she may have the same perceptual phenomenology that we do (needing no special mechanism to determine the orientation of that phenomenology, and so creating no risk of her becoming a partial zombie), but if we forced her post-perceptual systems into a perfectly symmetric format we would impair her cognition and so ensure that her cognitive phenomology (or top-down conceptual influences on her perceptual phenomenology) differed substantially from our own.[27]

[27]My thanks to David Anderson, Gregory Beauregard, David Chalmers, Brian Cutter, Dan Korman, Dana Goswick, Geoffrey Lee, Kristie Miller, John Morrison, Adam Pautz, Achille Varzi,

# References

Baars, B. J. (1988). A cognitive theory of consciousness. New York, NY: Cambridge University Press.

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.

Block, N. (2007). Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience. Behavioral and Brain Sciences 30 (5):481-548.

Brookshire, R. H. (2007). Introduction to neurogenic communication disorders (Seventh edition.). St. Louis, Mo.: Mosby Elsevier.

Budenz, D. (2008). Symmetry Between the Right and Left Eyes of the Normal Retinal Nerve Fiber Layer Measured with Optical Coherence Tomography (An AOS Thesis). Trans Am Ophthalmol Soc. 2008 Dec; 106: 252-275

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. (2019). Three puzzles about spatial experience. In A. Pautz & D. Stoljar (Eds.), Blockheads! Essays on Ned Block's philosophy of mind and consciousness (pp. 109-137). Cambridge: The MIT Press.

Guillemin, V. and Pollack, A. (1974). Differential Topology. Englewood Cliffs, N.J. : Prentice-Hall.

Hawthorne, J. (2004).Why Humeans Are Out of Their Minds. Noûs 38 (2):351-58.

Jennifer Wang, Douglas Zare, the judges of the 2018 Marc Sanders Prize in the Philosophy of Mind for insightful comments, and the editors of *Analytic Philosophy* for their patience and accommodation.

Hoefer, C. (2000). Kant's hands and Earman's pions: Chirality arguments for substantival space. International Studies in the Philosophy of Science 14:237-56.

Hurley, S. (1998). Consciousness in Action. Oxford University Press..

Kant, I. 1768. On the first ground of the distinction of regions in space. Translated in (J. Handyside, trans.) Kant's Inaugural Dissertation and Early Writings on Space, pp. 19-29. Open Court, 1929.

Lee, G. (2006). The experience of left and right. In (T. Gendler and J. Hawthorne, eds.) Perceptual Experience. Oxford University Press.

Lee, G. (2016). Synchronic Bertha.

Lycan, B. (2001). The Case for Phenomenal Externalism. Philosophical Perspectives 15 (15):17-35.

The LHCb collaboration: Aaij, R., Adeva, B. et al. (2015)é Determination of the quark coupling strength —Vub— using baryonic decays. Nature Phys 11, 743–747. https://doi.org/10.1038/nphys341

Maudlin, T. (1993). Buckets of Water and Waves of Space: Why Spacetime is Probably a Substance. Philosophy of Science 60(2): 183-203..

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.

Morrison, J. (2015). Anti-Atomism about Color Representation," Noûs.

Morrison, J. (2020). Perceptual Variation and Structuralism. Noûs.

Nerlich, G. (1979). Is Curvature Intrinsic to Physical Space? Philosophy of Science, 46: 439-58.

Nerlich, G. (2009). Incongruent Counterparts and the Reality of Space. Philosophy Compass 4/3: 598-613,

Noë, A. (2005). Action in Perception. MIT Press.

Pautz, A. (2017). The significance argument for the irreducibility of consciousness. Philosophical Perspectives 31 (1):349-407.

Pautz, A. (2019). How Can Brains in Vats Experience a Spatial World? A Puzzle for Internalists. In A. Pautz & D. Stoljar (Eds.), Blockheads! Essays on Ned Block's philosophy of mind and consciousness (pp. 109-137). Cambridge: The MIT Press

Perry, Z. (2017). How to be a Substantivalist without getting Shifty about it. Philosophical Issues, 27, Metaphysics. .

Pitt, D. (2004).

Pooley, O. (2003). Handedness, parity violation, and the reality of space. In (K. Brading & E. Castellani, eds) Symmetries in Physics: Philosophical Reflections. Cambridge University Press.

Rosenthal, D. (2005). Consciousness and Mind. Oxford University Press.

Shoemaker, S. (2006). Frege-Schlick.

Skow, B. (2007). Are Shapes Intrinsic? Philosophical Studies 133: 111-130.

Tye, M. (2000). Consciousness, Color and Content. MIT Press.

Volz, L. J., and Gazzaniga, M. S. (2017). Interaction in isolation: 50 years of insights from split-brain research. Brain 140, 2051-2060. doi: 10.1093/brain/awx139

Weatherson, B. (2005). Humeans Aren't Out of Their Minds. Noûs 41 (3):529-535.

# Appendix I

In this appendix I will elaborate on what I take to be the most promising option for the orientation-experience categoricalist. I will do three things. First, I will explain how to reject local supervenience without appealing to the fiction of lots of little gloves pervading the universe. Second, I will extoll the virtues of the resulting view: were it not for the problem of non-orientable worlds, this view should be the favorite. Third, I will argue that even given the problem of non-orientable worlds, we should stick with a version of this view if we want to maintain orientation-experience categoricalism, but the best way to do so

involves embracing property dualism (the denial that phenomenal properties are identical to or grounded in material properties).

## Part One

Recall the parable (introduced in §3.1) of the universe pervaded with gloves. In such a universe, we can divide neighborhoods of space or spacetime into equivalence classes based on the local congruency relations of the gloves. How might we extend this idea to worlds not pervaded with little gloves?

What is really doing the work is that you have some way of putting a series of points (or little bits of matter) in order, and then comparing their alignment while holding the order fixed. If the space in question is n-dimensional, then you only need as many points as it takes to determine an *ordered basis* for an n-dimensional vector space, which requires $n$ orthogonal vectors.[28]

A vector can be thought of as a line segment with both a length and a direction, which can be approximated by an ordered sequence of two points, the first being the origin and the second being the tip of the arrow. Thus an ordered basis for two-dimensional space can be approximated by four points of space or bits of matter (taken in a certain order), and an ordered basis for three-dimensional space can be approximated by six points (taken in a certain order), as long as none of the vectors so approximated are parallel to each other.[29]

In an actual vector space, there is a handy way to express the condition that two ordered bases are congruent: that the matrix expressing the linear transformation that maps one to the other (in either direction) has a positive determinant. Within an (abstract, mathematical) vector space, we do not have to relativize anything to a neighborhood: we can just ask of any two ordered bases for the space whether the matrix mapping one to the other has a positive determinant. This yields a partition of the set of ordered bases for the space into two equivalence classes. In the more general setting of manifold topology, we speak instead of local structure-preserving (homeomorphic) mappings called *coordinate charts*, which have overlapping domains and which each assign a system of (e.g. Cartesian) coordinates to a neighborhood (an open, compact region) in the space, and which together collectively map out the space (making something called an *atlas* of charts), with coherent structure-preserving transitions between systems of coordinates of charts whose domains overlap. To say that such an atlas can be specified (or approximated) on some physical space is to say that that space is locally approximated by a Euclidean space: a fairly minimal requirement that relationalists as well as substantivalists should hope to countenance about the actual physical world.[30]

[28] An ordered basis is an ordered sequence of vectors that mutually span the space, meaning that you can construct any vector as a combination of vectors in the basis, multiplied by scalars.

[29] We can also get by with just $n + 1$ points, by assuming that the first point serves as the origin for each vector.

[30] Some theories of quantum gravity suggest that spacetime is not continuous, which complicates

We can then define 'same orientation class as' (i.e. 'congruent with') relative to chart $c$ for the two-dimensional case as an 8-place relation, i.e. a comparison of two 4-tuples, whose elements are all in the domain of $c$, as follows:

SameOrientation$_c$($x_1$...$x_4$,$y_1$...$y_4$) iff the matrix transformation from ordered basis (c($x_1$)...c($x_4$)) to ordered basis (c($y_1$)...c($y_4$)) in $\mathbb{R}^2$ has a positive determinant.

This gives us a partition, restricted to the domain of $c$, of sequences (that is, of ordered tuples) of four unaligned points or small bits of matter into two equivalence classes, which we can call $A_c$ and $B_c$. In the case of a three-dimensional space the same technique works, applied to sequences of six elements rather than four.

Here, finally, is how we can use this to specify whether Righty experiences leftish or rightish. We first find some representative sequence of six elements within Righty's neural processing. In our toy example, this can be a measurement of his glove-shaped neural assembly. For example: first, a particle at the back of his palm, then a particle in the front of it (this measures the back-to-front-of-palm direction), then a particle at the base of his pointer finger, then a particle at its tip (this measures the base-to-tip-of-pointer-finger direction), then a particle at the base of this thumb and finally a particle in its tip (measuring the base-to-tip-of-outstretched-thumb direction: all of this assuming some canonical position, e.g. fingers parallel to arm, thumb maximally distant from palm). Next we (or, rather, the principles of psychophysics) choose a *phenomenal orientation*: which equivalence class in neighborhood $n$ mapped by chart $c$ gives rise to leftish experience, and which to rightish? Say class $A_c$ rather than $B_c$ gives rise to leftish. Then our account is that Righty has leftish experience in $n$, the domain of chart $c$, iff the representative sequence of his neural processing is c-congruent with elements of class $A_c$, and rightish experience iff it is c-congruent with elements of class $B_c$.

Now, this view may wind up committing us to a fairly extreme rejection of local supervenience. The picture, again, is that for Lefty to experience leftish is for his neural glove to be a member of one equivalence class rather than another: in effect, for a certain collection of SameOrientation$_c$ relations to obtain. However, we have seen that this cannot be the case at non-orientable worlds. This complicates things. Intuitively, you might think that the very same region $c$ and its inhabitants can exist at an orientable world like ours, and at a different world like Iris and Siri's that perfectly resembles ours around $c$ and differs only elsewhere.

But if that is correct, then say, the fact that Lefty's neural glove is congruent with elements of class A rather than class B cannot on its own be a sufficient ground for Lefty's leftish experience: we must add in some kind of global constraint on the topology, ensuring that the world in question is orientable.

To avoid this we might appeal to a theory like perdurantism, or sophisticated substantivalism, which will allow us to deny that the elements $x_1$...$x_4$,$y_1$...$y_4$ exist at more than

---

the story in ways I will ignore here.

one world (or more than one time at the same world). Then global topological constraints may play a systematizing role, but we can think of the specific world-bound entities within $c$ as being sufficient grounds for the experience facts, though their counterparts elsewhere will not always ground qualitatively identical experience facts. Our principle that locally congruent qualitative duplicates cannot differ experientially adds a further constraint, but it is of course a local one.[31]

## 5.3 Part Two

To sum up, the approach suggested here is a fully general one. It can be implemented no matter one's ontology, whether one is a relationist, a sophisticated substantivalist or a classical, 'manifold' substantivalist. Depending on the ontology of one's implementation, the view may have holistic implications. However, the underpinning of the view is still a series of local mappings (albeit subject to global constraints). This means that we can continue to think of the local features (i.e., local congruency relations) as being the *core* realizers of leftish and rightish (in Sydney Shoemaker 1984's sense), while the global configuration in light of which the global constraints are satisfied would be merely the non-core part of the *total* realizer.[32] Thus there is no commitment to action at a distance: anything that actually happens to Lefty's orientational experience (a change from leftish to rightish) can be explained in terms of goings on within his neighborhood.

In this subsection I will sing the praises of this account, setting aside the problem of non-orientable worlds. Recall spatial-orientation categoricalism, the naive view that there are intrinsic properties (that gloves can have) of left-handedness and right-handedness. Orientation-experience categoricalism fits with spatial-orientation categoricalism: pairing the two, we can say that the realizers of leftish experience use, or involve, instances of (intrinsic) left-handedness, and realizers of rightish experience use or involve instances of (intrinsic) right-handedness. Generally, realizers of orientational experience use the orientational properties that they are experiences of.

This compares with the way that we appear to use temporal properties in temporal experience, by way of temporal coding. Thus for example we use spike-timing dependent

[31]This choice between holism and world-bound individuals actually arises even before we consider Iris and Siri's world. If the relevant entities $x_1...x_4$,$y_1...y_4$ are endurant bits of matter rather than points of spacetime or world-bound, time-stamped perdurants, then they might all be in the leftish equivalence class at one time and then in the rightish equivalence class at a later time. Thus the realizers of leftish and rightish have to take into account where the relevant entities are in spacetime as a whole.

[32]We might expand to a three-way distinction, between core, local and distal components of the realizer, to stress the primal role of Lefty's neural state, the secondary role of local congruency relations in his neighborhood, and then the tertiary role of global background conditions. C.f. Wilson (2001).

coding to represent time-dependent acoustic signals, among other things.[33] In general, of course, it stands to reason that our brains would use spatiotemporal structural features to realize experience of them, given that our brains are spatiotemporally structured anyway.

Given that spatial-orientation categoricalism is false, the approach I've been considering above (let's call it the *local congruency approach*) is the closest we'll get to the ideal of our brains actually using orientational properties to realize experiences of them. Indeed, given that spatial-orientation categoricalism is false, orientational properties arguably *just are* the properties that emerge from the $\text{SameOrientation}_c$ relations.

This is not just an abstractly pleasing feature, it is what explains why the local congruency approach manages to deliver intuitive verdicts about what happens in otherwise difficult cases, cases involving symmetric experiencers, or experiencers who are left-right inverted with respect to one another.

We've already seen in §4 above that experiencers who are wired like Aviva are only a lesion away from becoming partial zombies. But this is only true if the realizers of Aviva's experience do not use actual orientational properties to realize her experience of them. If it is the actual orientation of her neural glove that determines whether she experiences rightish or leftish, irrespective of how it compares to some other internal landmark like the position of her Broca's area, then a lesion to her Broca's area needn't affect her leftish or rightish experience. Thus the local congruency view explains why Aviva is at no greater risk of becoming a zombie than the rest of us.

Finally, there is a secondary, puzzling implication of views that respect local supervenience, that the local congruency view avoids. Lee (2006) points out that if local supervenience is correct then a pair of mirror-reverse twins, say, Righty and Lefty, will be disposed to have mirror-opposite experiences of a given stimulus. That is, say that at $t_1$ Right and Lefty are genuine mirror-reverse twins. Local supervenience then implies that at $t_1$ they have the same experience. But presumably to maintain the perfect incongruent symmetry between them we have to maintain incongruent symmetry in their respective environments: if Righty is looking at his right hand, Lefty must be looking at his left hand. Say that the experience that both are having is leftish.

Now suppose that at $t_2$ we put them in the same environment, so that they are both looking at the same thing — say, Lefty's left hand. Now Lefty continues to have leftish experience, but Righty, having switched from looking at one kind of hand, to an incongruent counterpart of it, will switch from having leftish to having rightish experience. So Righty and Lefty are now looking at the same hand but Righty sees it as rightish while Lefty sees it as leftish (note that Righty and Lefty will no longer be perfect mirror-reverse twins at $t_2$: say, Righty will represent the hand as having its thumb pointed towards his Broca's area with palm facing back, Lefty as it having its thumb pointed away). In other words, Righty and Lefty are orientational experience spectrum inverts! This is puzzling for various reasons. For one thing we have no good way of assessing populations for this sort of

[33] See e.g. Di Lorenzo et. al. (2009).

inversion: given any group of people how uniform should we expect their orientational experience to be?

In contrast, on the local congruency approach, even at $t_1$ Righty and Lefty have different experiences, though they are genuine mirror-reverse twins. Righty can see his hand as rightish while Lefty can see his as leftish: there is no need to expect any degree of spectrum inversion. We can even make sense of how Righty's experience might reverse if he himself were to be reversed, say, by being flipped around through a fourth spatial dimension.[34]

## 5.4 Part Three

I conclude, then, that the local congruency approach is attractive, at least if we set aside the problem of non-orientable worlds. But what if we take that problem into account? It's complicated.

Supposing that we cannot bring ourselves to reject orientation-experience categoricalism, we must choose between taking local supervenience respecting views to be true at all possible worlds — ours and Aviva's and Iris and Siri's — or we must opt for a radical form of disjunctivism on which local supervenience respecting views are true on Iris and Siri's world (because this is necessary to respect the principle that locally congruent intrinsic duplicates are experiential duplicates) but local supervenience rejecting views like the local congruency approach are true on Aviva's world, and so presumably on ours (because this is necessary to keep Aviva from becoming a partial zombie).

Since it is the only way to both respect the principle that locally congruent intrinsic duplicates are experiential duplicates and keep partial zombies at bay (there may be nothing we can do for Aviva if she lives at Iris and Siri's world, but at least we can help her if she

[34]The details matter. Suppose you have him look at a left-handed glove, then ask him to close his eyes and sustain visual imagery of the glove ($t_1$). Then you quickly flip him around in a fourth spatial dimension ($t_2$). Then you ask him if anything has changed with his visual imagery ($t_3$, before he opens his eyes), then finally you ask him to open his eyes and say if anything has changed ($t_4$). Interestingly, all theories on the table predict the same behavior: at $t_3$ Righty will say that nothing has changed, and then at $t_4$ he will say that something has changed. But on the local congruency approach, there has been a change in Righty's visual imagery by $t_3$: his visual memory is now of a right-handed glove rather than a left-handed one, though Righty cannot access this change. Then at $t_4$ when he opens his eyes he will have a visual experience of a left-handed glove (as at $t_1$) but he will say that his experience has changed, because he can access the incongruency between his visual memory and his perceptual visual experience. Thus the local congruency approach is able to avoid the threat of pervasive mirror spectrum inverts, but at the cost of allowing that if someone is magically mirror-reversed, there will be certain changes to their phenomenal state which they cannot access. It is disputed whether consciousness normally or typically overflows access (Block 2007). But note that here, there is no *experience* which we claim overflows access. There is only a change between experiences that is inaccessible. We need not claim that this change itself is experienced. After all, the global workspace and metacognitive mechanisms will themselves be reversed in synch with the visual experience.

lives around here), the disjunctive approach seems to be our best bet. But is it tenable?

As I note in §5.1 above, the form of disjunctivism in question is far more radical than the functionalist's familiar sort of multiple realization. How can the same experience type — leftish experience of a glove, say — be realized by such a disparate collection of things? How can phenomenal similarity between the states of two things be grounded without an underlying metaphysical similarity in the realizers? how could Aviva's leftish be more similar to Iris' leftish than to, say, Aviva's rightish, or anything else, if the natures of their realizers are so radically different that one respects local supervenience and the other does not?

As I note in passing in §5.1, there is an adequate answer to this question: dualism. What we have to accept here is that the *material correlates* of the relevant experience are disjunctive, and unsuited to ground phenomenal similarity. Dualism (as I understand it here) is the view that phenomenal properties are neither identical to nor grounded in their material correlates. If leftish is a primitive (i.e. fundamental) property, we need no further explanation of the similarity of its instances. Then the disjunctivism of material correlates is an interesting consequence, but does not conflict with what we know about phenomenal similarity.

An added benefit of a dualist approach is that it mitigates the holistic implications of the local congruency view. Even if one accepts, say, an endurantist relationist metaphysics, so that global topological constraints figure in every material correlate of a leftish or rightish experience, still the experience itself might be a property that is intrinsic to its bearer, if dualism is true and phenomenal properties are not grounded in their material correlates. I have not suggested here that there is anything wrong with holism, but there are doubtless those who will think so.

Anyone opting for a dualist implementation must consider the question of causal efficacy. This is especially salient since, in §4 above, we argued that part of the problem with Aviva's being a partial zombie is that it deprives her leftish and rightish experiences of causal efficacy. Doesn't dualism do the same?

Two points in reply. First, the dualist can accommodate lack of causal efficacy (and the possibility of zombies) more naturally than the materialist can. For the dualist, zombies are an occupational hazard, and the reason to resist them is that they represent worlds with comparatively inelegant psychophysical laws (irrespective of causal efficacy). Second, there is only a deeply troubling problem of mental causation when candidate mental causes are undercut by clearly better suited, 'lower-level' physical causes. But all of the accounts on the table are sufficiently messy, that it is far from obvious that any such undercutting applies here. Maybe Righty's-neural-glove-being-in-local-equivalence-class-A has some claim to causal efficacy, but it is not obviously better suited than a fundamental, intrinsic rival, which is what the dualist says that a rightish phenomenal property would be.

I conclude that if we cannot bring ourselves to reject orientation-experience categoricalism — if we are sure that there is a categorical difference between leftish and rightish experience, much as there is a categorical difference between greenish and reddish expe-

rience — then our best bet is to opt for a dualistic form of disjunctivism, one according to which the material correlates of experience are given by the local congruency approach around here, and by a local supervenience respecting view at worlds like Iris and Siri's.